

(12)特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2004 年 11 月 18 日 (18.11.2004)

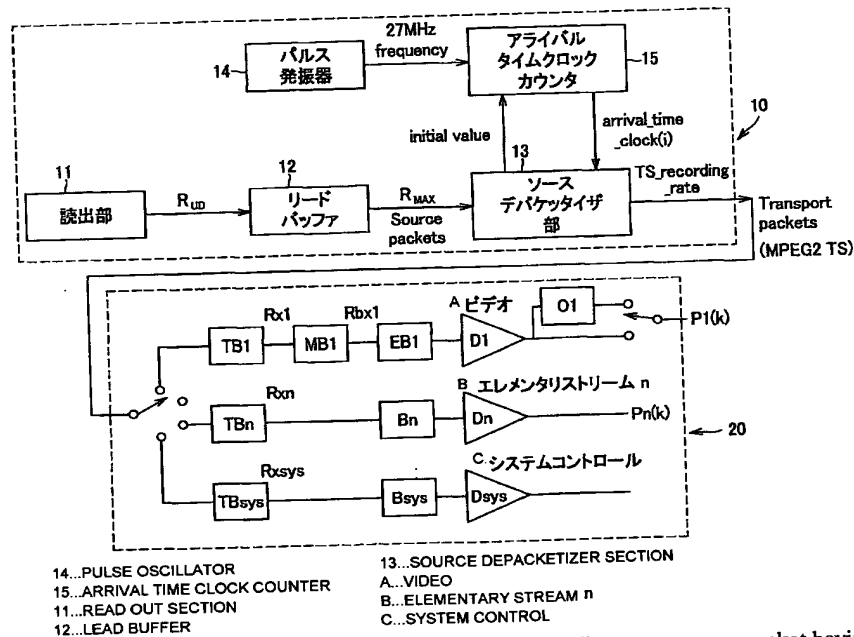
PCT

(10) 国際公開番号
WO 2004/100545 A1

- (51) 国際特許分類⁷: H04N 7/08, 7/24, 5/92
- (21) 国際出願番号: PCT/JP2004/005782
- (22) 国際出願日: 2004 年 4 月 22 日 (22.04.2004)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願2003-130661 2003 年 5 月 8 日 (08.05.2003) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 加藤 元樹 (KATO, Motoki) [JP/JP]; 〒1410001 東京都品川区北品川 6 丁目 7 番 3 5 号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 小池 晃, 外(KOIKE, Akira et al.); 〒1000011 東京都千代田区内幸町一丁目 1 番 7 号 大和生命ビル 11 階 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, ...)
- [続葉有]

(54) Title: INFORMATION PROCESSING DEVICE AND METHOD, PROGRAM, AND RECORDING MEDIUM

(54) 発明の名称: 情報処理装置及び方法, 並びにプログラム及び記録媒体



(57) Abstract: A player model (1) includes: an output section (10) for reading out a source packet having a transport packet and an arrival time stamp from TS1 and TS2 which are connected seamlessly and outputting the transport packet according to the arrival time stamp; and a decoder (20) for decoding the transport packet. The output section (10) inputs the transport packet to the decoder (20) according to the arrival time stamp of the source packet. The decoder (20) has an audio buffer TBn having a capacity capable of buffering the audio data corresponding to the time required for inputting the I picture as the first picture of TS2 into the video buffer TB1.

[続葉有]



NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

(57) 要約: プレーヤモデル(1)は、シームレスに接続されたTS1及びTS2から、トランスポートパケット及びアライバルタイムスタンプを有するソースパケットを読み出し、そのアライバルタイムスタンプに従ってトランスポートパケットを出力する出力部(10)と、トランスポートパケットをデコードするデコーダ(20)とから構成される。出力部(10)は、ソースパケットのアライバルタイムスタンプに従ってトランスポートパケットをデコーダ(20)に入力するものとし、デコーダ(20)は、そのオーディオバッファTBnを、TS2の最初のピクチャであるIピクチャをビデオバッファTB1に入力するために要する時間分のオーディオデータをバッファリング可能な容量とする。

明細書

情報処理装置及び方法、並びにプログラム及び記録媒体

技術分野

本発明は、ビデオストリームとオーディオストリームとの多重化ストリームをビデオフレーム精度で編集して、編集点を連続的（シームレス）に再生するための情報処理装置、その方法、プログラム及び記録媒体、並びにシームレス再生に最適な形式で多重化ストリームを生成する情報処理装置及びその多重化されたストリームデータが記録された記録媒体に関する。

本出願は、日本国において2003年5月8日に出願された日本特許出願番号2003-130661を基礎として優先権を主張するものであり、この出願は参照することにより、本出願に援用される。

背景技術

ビデオストリームとオーディオストリームの多重化ストリームをビデオフレーム精度で編集して、編集点をシームレスに再生するための方法が例えば日本公開特許公報2000-175152号、日本公開特許公報2001-544118号、日本公開特許公報2002-158974号に記載されている。

図1は、従来のDVR-STDモデル（DVR MPEG2 transport stream player model）（以下、プレーヤという。）101を示すブロック図である。DVR-STDは、シームレス接続された2つのPlayItemによって参照されるAVストリームの生成及び検証の際におけるデコード処理をモデル化するための概念モデルである。

図1に示すように、プレーヤ101においては、読出部（DVRdrive）111からビットレート R_{ud} で読み出されたTS（トランスポートストリーム）ファイルは、リードバッファ112にバッファリングされ、このリードバッファ112からソースパケットがソースデパケッタイザ部（source depacketizer）113へ最大ピ

ットレート R_{MAX} で読み出される。

パルス発振器 (27MHz X-tal) 114は、27MHzのパルスを発生する。アライバルクロックカウンタ (Arrival time clock counter) 115は、この27MHzの周波数のパルスをカウントするバイナリカウンタであり、ソースデパケッタイザ部113に、時刻 $t(i)$ におけるArrival time clock counterのカウント値Arrival_time_clock(i)を供給する。

1つのソースパケットは、1つのトランスポートパケットとそのarrival_time_stampを持つ。現在のソースパケットのarrival_time_stampがarrival_time_clock(i)のLSB (least significant Bit: 最下位ビット) 30ビットの値と等しいとき、ソースデパケッタイザ部113からそのトランスポートパケットが出力される。TS_recording_rateは、トランスポートストリーム (以下、TSという。) のビットレートである。また、図12に示す n 、 TB_n 、 MB_n 、 EB_n 、 TB_{sys} 、 B_{sys} 、 R_{xn} 、 R_{bxn} 、 R_{xsys} 、 D_n 、 D_{sys} 、 O_n 及び $P_n(k)$ の表記方法は、ISO/IEC13818-1

(MPEG2 Systems規格) のT-STD (ISO/IEC 13818-1で規定されるtransport stream system target decoder) に定義されているものと同じである。

次に、このような従来のプレーヤ101におけるデコーディングプロセスについて説明する。まず、1つのDVR MPEG2 TSを再生しているときのデコーディングプロセスを説明する。単一のDVR MPEG2 TSを再生している間は、出力部110からトランスポートパケットを出力してデコーダ120であるDVR-STDの TB_1 、 TB_n 又は TB_{sys} のバッファへ入力するタイミングは、ソースパケットのarrival_time_stampにより決定される。 TB_1 、 MB_1 、 EB_1 、 TB_n 、 B_n 、 TB_{sys} 及び TB_{sys} のバッファリング動作の規定は、ISO/IEC 13818-1に規定されているT-STDと同じである。復号動作と表示動作の規定もまた、ISO/IEC 13818-1に規定されているT-STDと同じである。

次に、シームレス接続されたPlayItemを再生している間のデコーディングプロセスについて説明する。ここでは、シームレス接続されたPlayItemによって参照される先行するストリームTS1と、現在のストリームTS2との再生について説明をする。

あるAVストリーム (TS1) からそれにシームレスに接続された次のAVストリ

ーム (TS2) へと移る間には、TS2のアライバルタイムベースの時間軸は、TS1のアライバルタイムベースの時間軸と同じ時間軸でない。また、TS2のシステムタイムベースの時間軸は、TS1のシステムタイムベースの時間軸と同じ時間軸でない。ビデオの表示は、シームレスに連続していることが要求される。オーディオのプレゼンテーションユニットの表示時間にはオーバーラップがあってもよい。

次に、ソースデパケッタイザ部から読み出されるトランスポートパケットのDVR-STDへの入力タイミングについて説明する。

(1) TS1の最後のビデオパケットがDVR-STDのTB1に入力終了する時刻 T_1 までの時間

時刻 T_1 までの時間は、DVR-STDのTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパケットのarrival_time_stampによって決定される。

(2) 時刻 T_1 からTS1の残りのパケットの最後のバイトが入力終了する時刻 T_2 まで

TS1の残りのパケットは、TS_recording_rate (TS1) のビットレート (TS1の最大ビットレート) でDVR-STDのTBn又はTBsysのバッファへ入力されなければならない。ここで、TS_recording_rate (TS1) は、Clip1に対応するClipInfo()において定義されるTS_recording_rateの値である。TS1の最後のバイトがバッファへ入力する時刻は、時刻 T_2 である。したがって、時刻 T_1 から T_2 までの区間では、ソースパケットのarrival_time_stampは無視される。

N_1 をTS1の最後のビデオパケットに続くTS1のトランスポートパケットのバイト数とすると、時刻 T_1 から T_2 までの時間 $T_{2-1} = T_2 - T_1$ は、 N_1 バイトがTS_recording_rate (TS1) のビットレートで入力終了するために必要な時間であり、下記式 (1) のように計算される。

$$T_{2-1} = T_2 - T_1 = N_1 / \text{TS_recording_rate (TS1)} \quad \dots (1)$$

時刻 T_1 から T_2 までの間は、図1に示すRxnとRsysの値は共にTS_recording_rate (TS1) の値に変化する。このルール以外のバッファリング動作は、T-STDと

同じである。

オーディオデコーダは、時刻 T_1 から T_2 までの区間の入力データを処理することができるように、即ち、時刻 T_1 から T_2 までの間は、図 1 に示す R_{xn} と R_{sys} の値が共に $TS_recording_rate$ ($TS1$) の値に変化するため、T-STD で定義されるバッファ量に加えて付加的なバッファ量（約 1 秒分のデータ量）が必要である。

(3) 時刻 T_2 以後

T_2 の時刻において、アライバルタイムクロックカウンタ 115 は、 $TS2$ の最初のソースパケットの $arrival_time_stamp$ の値にリセットされる。DVR-STD の $TB1$ 、 TBn 又は TB_{sys} のバッファへの入力タイミングは、 $TS2$ のソースパケットの $arrival_time_stamp$ によって決定される。 R_{xn} と R_{sys} はともに、T-STD において定義されている値に変化する。

次に、ビデオのプレゼンテーションタイミングについて説明する。ビデオプレゼンテーションユニットの表示は、接続点（コネクションポイント）を通してギャップなしに連続でなければならない。

ここで、

STC (System Time Clock) 1 : $TS1$ のシステムタイムベースの時間軸

STC2 : $TS2$ のシステムタイムベースの時間軸（正確には、STC2 は、 $TS2$ の最初の PCR (Program Clock Reference) が T-STD に入力した時刻から開始する。）

とする。

STC1 と STC2 との間のオフセット値は、次のように決定される。

$PTS1_{end}$: $TS1$ の最後のビデオプレゼンテーションユニットに対応する STC1 上の PTS

$PTS2_{start}$: $TS2$ の最初のビデオプレゼンテーションユニットに対応する STC2 上の PTS

T_{pp} : 最後のビデオプレゼンテーションユニットの表示期間

とすると、2 つのシステムタイムベースの間のオフセット値 STC_delta は、下記式

(2) のように計算される。

$$STC_delta = PTS1_{end} + T_{pp} - PTS2_{start} \quad \dots (2)$$

次に、オーディオのプレゼンテーションタイミングについて説明する。TS1とTS2との接続点において、オーディオプレゼンテーションユニットの表示タイミングのオーバーラップがあってもよく、それは0から2オーディオフレーム未満である。プレーヤ101には、どちらのオーディオサンプルを選択するかということと、オーディオプレゼンテーションユニットの表示を接続点の後の補正されたタイムベースに再同期する制御が要求される。

TS1からそれにシームレスに接続された次のTS2へと移るとき、DVR-STDのシステムタイムクロックの制御について、プレーヤ101が行う処理を説明する。TS1の最後のオーディオプレゼンテーションユニットが表示される時刻T5において、システムタイムクロックは、時刻T2からT5の間にオーバーラップしていてもよい。この区間では、DVR-STDは、システムタイムクロックを古いタイムベースの値(STC1)から新しいタイムベースの値(STC2)の値に切り替える。STC2の値は、下記式(3)のように計算できる。

$$STC2 = STC1 - STC_delta \quad \dots (3)$$

TS1から、これにシームレスに接続された次のTS2へと移るとき、TS1及びTS2が満たさなければいけない符号化条件を説明する。

STC1¹_{video_end} : TS1の最後のビデオパケットの最後のバイトがDVR-STDのTB1へ到着するときのシステムタイムベースSTC1上のSTCの値

STC2²_{video_start} : TS2の最初のビデオパケットの最初のバイトがDVR-STDのTB1へ到着するときのシステムタイムベースSTC2上のSTCの値

STC2¹_{video_end} : STC1¹_{video_end}の値をシステムタイムベースSTC2上の値に換算した値

とした場合、STC2¹_{video_end}は、下記式(4)のようにして計算される。

$$STC2^{1_video_end} = STC1^{1_video_end} - STC_delta \quad \dots (4)$$

ここで、デコーダ 1 2 0 が DVR-STD に従うために、次の 2 つの条件を満たすことが要求される。

(条件 1)

TS2 の最初のビデオパケットの TB1 への到着タイミングは、次に示す不等式 (5) を満たさなければならない。

$$STC2^{2_video_start} > STC2^{1_video_end} + T_{2-1} \quad \dots (5)$$

この上記不等式 (5) を満たすように、Clip 1 及び／又は Clip 2 の部分的なストリームを再エンコード及び／又は再多重化することが必要になる。

(条件 2)

STC1 と STC2 とを同じ時間軸上に換算したシステムタイムベースの時間軸上において、TS1 からのビデオパケットの入力とそれに続く TS2 からのビデオパケットの入力は、ビデオバッファをオーバーフロー及びアンダフローさせてはならない。

しかしながら、上述したように、DVR-STD モデルを使用した従来のプレーヤ 1 0 1 においては、時刻 T 1 から T 2 までの区間の入力データを処理することができるよう、即ち、時刻 T 1 から T 2 までの間は、TS1 の残りのパケットは、TS_recording_rate (TS1) のビットレート (TS1 の最大ビットレート) で DVR-STD の TBn 又は TBsys のバッファへ入力されていたため、T-STD で定義されるバッファに加えて、約 1 秒分のデータをバッファリング可能な容量の付加的なバッファが必要である。

このバッファ容量の大きさは、次の要因に基づく。即ち、MPEG2TS の中において、あるバイト位置にあるビデオデータに同期再生されるオーディオデータが、所定範囲内の多重化位相差を離れて存在することができ、この多重化位相差の最大値が 1 秒分のデータ量に相当する。したがって、上記式 (1) の N 1 の最大値は、最大 1 秒分のオーディオデータに相当する。時刻 T 1 から T 2 までの間は、ソー

スパケットのarrival_time_stampを無視して、TSの最大ビットレートでN1のデータ量のソースパケットが、オーディオバッファに入力されるので、このデータ量をバッファリングするために、T-STDで定義されるバッファ量に加えて付加的なバッファ量（約1秒分のデータ量）が必要になっていた。

この付加的なバッファの大きさを具体的に計算すると次のようになる。即ち、例えば640 kbpsのドルビーAC3方式により符号化されたオーディオストリームの場合、1秒分のオーディオデータは、640 kbits=80 kBytesとなり、80 kBytesの付加的なバッファが必要となる。

また、Linear PCM方式により符号化されたオーディオストリームであって、24 bitSample、96 kHz sampling frequency、8 channelsの場合、1秒分のオーディオデータは、 $24 \text{ bitSample} \times 96000 \text{ samples/sec} \times 8 \text{ channels} = \text{約 } 18 \text{ Mbits}$ となり、約3 Mbytesの付加的なバッファが必要となり、このようなマルチチャンネルのオーディオデータを扱う場合、この付加的なバッファが極めて大きくなってしまいうという問題点がある。

発明の開示

本発明は、このような従来の実情に鑑みて提案されたものであり、オーディオストリームとビデオストリームとが多重化された2つの多重化ストリームをシームレスに連続して復号するために、最適な容量のオーディオバッファとした情報処理装置、その方法、プログラム及び記録媒体、並びにこのようなオーディオバッファ容量に対応した多重化ストリームを生成する情報処理装置及びその方法、並びにその多重化ストリームが記録された記録媒体を提供することを目的とする。

上述した目的を達成するために、本発明に係る情報処理装置は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理装置において、上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソー

スパケットを出力する出力手段と、上記ソーススパケットのうちビデオデータをバッファリングするビデオバッファと、上記ソーススパケットのうちオーディオデータをバッファリングするオーディオバッファと、上記ビデオバッファにバッファリングされたビデオデータを復号するビデオ復号手段と、上記オーディオバッファにバッファリングされたオーディオデータを復号するオーディオ復号手段とを有し、上記オーディオバッファは、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量を有することを特徴とする。

本発明においては、オーディオバッファのサイズを、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量とするとともに、上記第1のピクチャが上記ビデオバッファに入力終了する時刻から第1の多重化ストリームの最後のソーススパケットを入力終了する時刻までの間においても、多重化ストリームのソーススパケットのアライバルタイムスタンプ (arrival_time_stamp) に従ってソーススパケットをバッファに入力するようにしたので、従来ソーススパケットのarrival_time_stampを無視して、トランスポートストリーム (TS) の最大ビットレートで入力するために必要とされていた1秒分の付加的なバッファを不要とするとともに、第1の多重化ストリームの最後のトランスポートパケットを入力した後に、第2の多重化ストリームの最初にデコードされるピクチャをそのデコードタイミングまでにビデオバッファへ入力することができる。

また、上記オーディオバッファの必要な容量を EBn_max (bits) とし、上記第2のピクチャのビット量を I_max (bits) とし、上記ビデオバッファへの入力ビットレートを Rv (bps) とし、オーディオデータのビットレートを Ra (bps) としたとき、 $EBn_max = (I_max / Rv) \times Ra$ を満たすものとしてことができ、第2のピクチャとなる例えばIピクチャのビット量の最大値を I_max とした場合、オーディオバッファの容量は、最大で $EBn_max = (I_max / Rv) \times Ra$ とすることができる。

更に、上記オーディオバッファは、100ミリ秒分のオーディオデータをバッファリング可能な容量を有することが好ましく、これにより、例えばMPEG2のIピクチャのデータ量は、通常1秒間で伝送されるデータ量の10%以下の大きさで

あるため、オーディオバッファをこれと同じ大きさの容量としておき、その大きさ分のオーディオデータを先送りしておくことで、Iピクチャをそのデコードタイミングまでにビデオバッファへ入力することができ、ビデオデータのエンコード制限が少なくなる。即ち、オーディオバッファを上記容量とすることで、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く入力終了するように多重化ストリームを多重化することができる。

更にまた、上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値 $STC1^{1_end}$ を上記時間差STC_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値を $STC2^{1_end}$ ($=STC1^{1_end} - STC_delta$) とし、上記第2の多重化ストリームの最初のソースパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値を $STC2^{2_start}$ としたとき、上記多重化ストリームは、 $STC2^{2_start} > STC2^{1_end}$ を満たすものとするにより、DVR-STDに従ったものとすることができる。

更にまた、上記第1の多重化ストリームの最後のソースパケットが上記出力手段から出力された後、所定時間delta1の経過後、上記第2の多重化ストリームの最初のソースパケットを上記出力手段から出力するものとし、 $STC2^{2_start} > STC2^{1_end} + delta1$ を満たすものとしてもよく、これにより、第2の多重化ストリームの最初のソースパケットの入力タイミングに柔軟性を持たせ、第2の多重化ストリームの符号化を容易にすることができる。

更にまた、上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC_deltaとし、上記第1の多重化ストリームの最後のソースパケットの出力を開始した後、所定時間ATC_deltaの経過後に上記第2の多重化ストリームの最初のソースパケットを上記出力手段から出力するものとし、上記所定時間ATC_deltaを、上記時間差STC_deltaの値を満たすように決定し、上記多重化ストリームは、上記時間差STC_deltaの値を満たすように多重化さ

れたものとしてもよく、これにより、第2の多重化ストリームの最初のソースパケットの入力タイミングに柔軟性を持たせ、第2の多重化ストリームの符号化を容易にすることができる。

このとき、上記所定時間ATC_deltaの値は、上記第1の多重化ストリームの付属情報として管理することができる。

本発明に係る情報処理方法は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理方法において、上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、上記ビデオバッファ及びオーディオバッファにバッファリングされたビデオデータ及びオーディオデータを復号する復号工程とを有し、上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングすることを特徴とする。

また、本発明に係るプログラムは、上述した情報処理をコンピュータに実行させるものであり、本発明に係る記録媒体は、そのようなプログラムが記録されたコンピュータ読取可能なものである。

本発明に係る他の記録媒体は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなる多重化ストリームが記録された記録媒体であって、上記多重化ストリームは、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続され、上記第1及び第2の多重化ストリームが夫々のアライバルタイムスタンプに基づいてデコーダに入力可能であって、且つ、上記第2のピクチャをデコーダに入力する

ために要する時間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録されたことを特徴とする。

本発明においては、第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化ストリームが多重化されているため、このような多重化ストリームを、第2のピクチャをビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量のオーディオバッファを有するデコーダにより復号すれば、第1の多重化ストリームの最後のトランスポートパケットを入力した後に、第2の多重化ストリームの最初にデコードされるピクチャをそのデコードタイミングまでにビデオバッファへ入力することができる。

本発明に係る他の情報処理装置は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理装置において、第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化手段と、上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化手段とを有し、上記多重化手段は、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化することを特徴とする。

本発明においては、上記第2のピクチャを上記デコーダに入力するために要する時間分の、例えば100ミリ秒分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化するため、デコーダにおいて、オーディオデータをオーディオバッファに先送りして、1ピクチャ等の第2のピクチャを、そのデコードタイミングまでに伝送する時間を十分に確保することができ、多重化ストリームの符号化が容易になる。

本発明に係る他の情報処理方法は、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理方法において、第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化工程と、上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化工程とを有し、上記多重化工程では、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化することを特徴とする。

本発明の更に他の目的、本発明によって得られる具体的な利点は、以下に説明される実施例の説明から一層明らかにされるであろう。

図面の簡単な説明

図1は、従来の情報処理装置を示すブロック図である。

図 2 は、Bridge-Clipを使用する場合の先行するPlayItemと現在のPlayItemとの関係を示す模式図である。

図 3 は、Bridge-Clipを使用しない場合の先行するPlayItemと現在のPlayItemの関係を示す模式図である。

図 4 は、ビデオストリームとしてのClip1とClip2とをシームレス接続する例をピクチャの表示順序 (Presentation order) で示す模式図である。

図 5 は、図 4 に表示ビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、第 1 の方法であるBridgeSequenceを使用してシームレス接続を実現する各 A V ストリームにおけるデータ列を示す模式図である。

図 6 は、図 4 示すビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、第 2 の方法であるBridgeSequenceを使用しないでシームレス接続を実現する各 A V ストリームにおけるデータ列を示す模式図である。

図 7 は、オーディオの表示のオーラップを説明する図であって、TS1及びTS2におけるビデオのプレゼンテーションユニット及びオーディオのプレゼンテーションユニットを示す模式図である。

図 8 は、本発明の実施の形態における情報処理装置を示すブロック図である。

図 9 は、ある A V ストリーム (TS1) から、これにシームレスに接続された次の A V ストリーム (TS2) へと移るときのトランスポートパケットの入力、復号、及び表示のタイミングチャートである。

図 10 は、ある A V ストリーム (TS1) からそれにシームレスに接続された次の A V ストリーム (TS2) へと移るときのトランスポートパケットの入力、復号、表示の他の例を示すタイミングチャートである。

図 11 は、ある A V ストリーム (TS1) からそれにシームレスに接続された次の A V ストリーム (TS2) へと移るときのトランスポートパケットの入力、復号、表示の他の例を示すタイミングチャートである。

図 12 は、ATC_deltaを格納するための付属情報ClipInfo() のデータフォーマットを示す図である。

図 13 は、ある A V ストリーム (TS1) に接続されるところの次の A V ストリーム (TS2) が複数存在する場合の付属情報ClipInfo() を示す模式図である。

図 1 4 は、従来の DVR-STD の場合であり、オーディオのバッファサイズが 4 kBytes である場合に、TS1 からそれにシームレスに接続される次の TS2 へと移るときの DVR-STD のビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。

図 1 5 は、本発明の実施の形態における効果を説明する図であって、オーディオのバッファサイズが 8 kBytes である場合に、TS1 からそれにシームレスに接続される次の TS2 へと移るときの DVR-STD のビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。

発明を実施するための最良の形態

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、ビデオストリームとオーディオストリームとが多重化された 2 つの AV ストリームをシームレスで連続再生する情報処理装置に適用したものである。そして、本実施の形態においては、DVR-STD (Digital Video Recording-System Target Decoder) において、シームレスに接続された 2 つの AV ストリームを再生する際に最適な容量としたオーディオバッファを提案するものである。

はじめに、以下の説明において使用する用語について説明する。Clip は、ビデオストリームとオーディオストリームとの多重化ストリームを示す。また、Play List は、Clip 中の再生区間の集まりを示す。ある Clip 中の 1 つの再生区間は、PlayItem と呼ばれ、それは、時間軸上の IN 点と OUT 点のペアで表される。それゆえ、PlayList は、PlayItem の集まりである。

PlayItem 間をシームレスに再生するとは、再生装置（プレーヤ）が、デコーダの再生出力にポーズやギャップを起こさせることなく、ディスクに記録されたオーディオビデオデータを表示（再生）することができることである。

シームレス接続されている 2 つの PlayItems の構造について説明する。先行する PlayItem と現在の PlayItem との接続が、シームレス表示できるように保証されているかどうかは、現在の PlayItem において定義されている connection_condition

フィールドから判断することができる。そして、PlayItem間のシームレス接続は、Bridge-Clip (BridgeSequence) を使用する方法（以下、第1の方法とする。）と使用しない方法（以下、第2の方法とする。）がある。

まず、先行するPlayItem (previous PlayItem) と現在のPlayItem (current Play Item) とが、第1の方法であるBridgeSequenceを使用して接続されている場合のTS1及びTS2について説明する。図2は、第1の方法であるBridge-Clipを使用する場合における、先行するPlayItemであるP I 1と現在のPlayItemであるP I 2との関係を示す模式図である。この図2において、Bridge-Clipを使用する場合にプレーヤが読み出すストリームデータを影を付けて示す。ここで、DVR MPEG

(Moving Picture Experts Group) 2トランスポートストリーム (TS) は、整数個のAligned unitから構成される。Aligned unitの大きさは、6144バイト (2048×3バイト) である。1つのAligned unitは、32個のソースパケットからなり、ソースパケットの第1バイト目から始まる。

各ソースパケットは、192バイト長である。1つのソースパケットは、TP_extra_headerとトランスポートパケットとからなり、TP_extra_headerは、4バイト長であり、またトランスポートパケットは、188バイト長である。TP_extra_headerは、copy_permission_indicatorとarrival_time_stampとを有し、copy_permission_indicatorは、トランスポートパケットのペイロード (Payload) のコピー制限を表す整数、arrival_time_stamp (ATS) は、AVストリームの中で、対応するトランスポートパケットがデコーダに到着する時刻を示すタイムスタンプである。AVストリームを構成する各ソースパケットのarrival_time_stampに基づいて作られる時間軸をアライバルタイムベースといい、そのクロックをATC (Arrival Time Clock) と呼ぶ。

図2に示すTS1 (第1の多重化ストリーム) は、Clip1 (Clip AVストリーム) の影を付けて示すストリームデータD1とBridge-ClipのSPN_arrival_time_discontinuityより前の影を付けて示すストリームデータD2からなる。SPN_arrival_time_discontinuityは、the Bridge-Clip AVストリームファイルの中でアライバルタイムベースの不連続点があるところのソースパケットのアドレスを示す。

そして、TS1に含まれるClip1の影を付けて示すストリームデータD1は、先行

するPlayItemのIN_time（図2においてIN_time 1で示す。）に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスから、SPN_exit_from_previous_Clipで参照されるソースパケットまでのストリームデータである。

また、TS1に含まれるBridge-ClipのSPN_arrival_time_discontinuityより前の影を付けて示すストリームデータD 2は、Bridge-Clipの最初のソースパケットから、SPN_arrival_time_discontinuityで参照されるソースパケットの直前のソースパケットまでのストリームデータである。

また、図2に示すTS2（第2の多重化ストリーム）は、Clip 2（Clip AVストリーム）の影を付けて示すストリームデータD 4とBridge-ClipのSPN_arrival_time_discontinuity以後の影を付けて示すストリームデータD 3からなる。

そして、TS2に含まれるBridge-ClipのSPN_arrival_time_discontinuity以後の影を付けて示すストリームデータD 3は、SPN_arrival_time_discontinuityで参照されるソースパケットから、Bridge-Clipの最後のソースパケットまでのストリームデータである。

また、TS2に含まれるClip 2の影を付けて示すストリームデータD 4は、SPN_enter_to_current_Clipで参照されるソースパケットから、現在のPlayItemのOUT_time（図2において、OUT_time 2で示す。）に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスまでのストリームデータである。

次に、先行するPlayItemと現在のPlayItemが、第2の方法であるBridgeSequenceを使用しないで接続されている場合のTS1及びTS2について説明する。図3は、第2の方法であるBridge-Clipを使用しない場合における、先行するPlayItemであるP I 1と現在のPlayItemであるP I 2との関係を示す模式図である。この図3において、プレーヤが読み出すストリームデータは、影を付けて示す。

図3に示すTS1（第1の多重化ストリーム）は、Clip 1（Clip AVストリーム）の影を付けて示すストリームデータD 5からなる。TS1に含まれるClip 1の影を付けて示すストリームデータD 5は、先行するPlayItemのIN_time（図3においてIN_time 1で示す）に対応するプレゼンテーションユニットを復号するために必要

なストリームのアドレスから始まり、Clip 1 の最後のソースパケットまでのデータである。

また、図 3 に示す TS2 (第 2 の多重化ストリーム) は、Clip 2 (Clip AV ストリーム) の影を付けて示すストリームデータ D 6 からなる。TS2 に含まれる Clip 2 の影を付けて示すストリームデータ D 6 は、Clip 2 の最初のソースパケットから始まり、現在の PlayItem の OUT_time (図 3 において OUT_time 2 で示す) に対応するプレゼンテーションユニットを復号するために必要なストリームのアドレスまでのストリームデータである。

図 2 及び図 3 の両方において、TS1 及び TS2 は、ソースパケットの連続したストリームである。次に、TS1 及び TS2 のストリーム規定とそれらの間の接続条件について説明する。

TS1 及び TS2 は、ビデオストリームとオーディオストリームとが多重化されたものであるが、ここでは、先ず、シームレス接続のための符号化制限におけるビデオビットストリームの制限について説明する。

図 4 は、ビデオストリームとしての Clip1 と Clip2 とをシームレス接続する例をピクチャの表示順序 (Presentation order) で示す模式図である。動画プログラム的一部分をスキップ再生する際に、スキップ再生開始点であるアウト点ピクチャより時間的に前側のプログラムであるアウト点側プログラムと、スキップ再生到達点であるイン点ピクチャより時間的に後側のプログラムであるイン点側プログラムとをシームレスに接続するには、復号装置においてビデオストリームの再エンコード処理を行う必要がある。

MPEG2 規格に準じた画像群の単位である GOP (group of pictures) には、他の画像からの予測符号化なしに画像が符号化された参照画像である少なくとも 1 つの I (Intra) ピクチャ (フレーム内符号化画像) と、表示順序に順方向の予測符号化を用いて画像が符号化された順方向予測符号化画像である P (predictive) ピクチャと、順方向及び逆方向の予測符号化を用いて画像が符号化された双方向予測符号化画像である B (bidirectionally) ピクチャとの 3 種類の符号化画像が含まれている。図 4 においては、Clip1 及び Clip2 の各数字は表示順序を示し、I、P、B、又は i、p、b は各ピクチャの種類を示す。例えば図 4 には、Clip

1のB 7とClip 2のb 4とを接続する例を示すが、この接続点においてビデオストリームをシームレスに表示できるためには、OUT_time1 (Clip1のOUT_time) の後とIN_time2 (Clip2のIN_time) の前に表示される不必要なピクチャは、接続点付近のClipの部分的なストリームを再エンコードするプロセスにより、除去されなければならない。

図4に示すようなビデオストリーム (Clip1及びClip2) をシームレス接続する場合に、上記第1の方法であるBridgeSequenceを使用してシームレス接続を実現する例を図5に示す。SPN_arrival_time_discontinuityより前のBridge-Clipのビデオストリームは、図4に示すClip1のOUT_time1に対応するピクチャまでの符号化ビデオストリームからなる。そして、そのビデオストリームは先行するClip1のビデオストリームに接続され、1つの連続でMPEG2規格に従ったエレメンタリストリームとなるように再エンコードされている。同様に、SPN_arrival_time_discontinuity以後のBridge-Clipのビデオストリームは、図4のClip2のIN_time2に対応するピクチャ以後の符号化ビデオストリームからなる。そして、そのビデオストリームは、正しくデコード開始することができて、これに続くClip2のビデオストリームに接続され、1つの連続でMPEG2規格に従ったエレメンタリストリームとなるように再エンコードされている。Bridge-Clipを作るためには、一般に、数枚のピクチャは再エンコードしなければならない、それ以外のピクチャはオリジナルのClipからコピーすることができる。

図5の(a)は、図4に示すClip1をその復号順に示したものであり、プレーヤーは、先行するClip1のP 5のソースパケット番号 (SPN_exit_from_previous_Clip) から(b)に示すBridge-Clipにジャンプする。ここで、Bridge-Clipの図2に示したD 2、即ち、BridgeSequenceにおけるSPN_arrival_time_discontinuityの前までのビデオデータに相当するClip1のOUT_time1側のストリームデータにおいて、B 4までのデータd 1は、Clip1をそのままコピーしたデータからなり、これに続くデータP 7、B 6は本来はオリジナルのClip1のB 6及びB 7からなるが、Clip1を復号し圧縮していない画像データに戻してから再びエンコードしてP 7及びB 6としたデータd 2となっている。また、Bridge-Clipの図2に示したD 3、即ち、BridgeSequenceにおけるSPN_arrival_time_discontinuity以降のビデオデ

ータに相当するClip2のIN_time2側のストリームデータにおいても、オリジナルのClip2のb 4、p 5、p 8、b 6、b 7は、Clip2を一旦復号して圧縮していない画像データに戻してから再びエンコードして新たに作成されたデータ（i 0、p 1、p 4、b 2、b 3）d 3となり、それ以降Clip2のSPN_enter_to_currenr_Clipにジャンプするまでの間のデータd 4は、Clip2をそのままコピーしたものとなっている。

次に、図4に示すようなビデオストリーム（Clip1及びClip2）をシームレス接続する場合に、上記第2の方法であるBridgeSequenceを使用しないでシームレス接続を実現する例を図6に示す。図6において、Clip1及びClip2は、その復号順序にピクチャを示したものである。BridgeSequenceを使用しない場合であっても、図5に示したBridgeSequenceを使用する場合と同様に、接続点（コネクションポイント：conection point）付近のストリームは、一旦復号して圧縮していないデータに戻してから最適なピクチャタイプに再デコードされる。即ち、Clip1のビデオストリームは、図4のOUT_time1に対応するピクチャまでの符号化ビデオストリームからなり、それは、1つの連続でMPEG2規格に従ったエレメンタリストリームとなるように、オリジナルのClip1のB 6、B 7が再エンコードされたデータ（P 7、B 6）d 5とされている。同様にして、Clip2のビデオストリームは、図4のClip2のIN_time2に対応するピクチャ以後の符号化ビデオストリームからなり、それは、1つの連続でMPEG2規格に従ったエレメンタリストリームとなるように、オリジナルのClip2のb 4、p 5、p 8、b 6、b 7が再エンコードされたデータ（i 0、p 1、p 4、b 2、b 3）d 6とされている。

次に、TS1及びTS2のそれぞれの多重化ストリームの符号化制限について説明する。図7は、オーディオの表示のオーバーラップを説明する図であって、TS1及びTS2におけるビデオのプレゼンテーションユニットVPU1、VPU2及びオーディオのプレゼンテーションユニットAPU1、APU2を示す模式図である。

図7に示すように、TS1のオーディオストリームの最後のオーディオフレームA_endは、TS1の最後の表示ピクチャの表示終了時（OUT_time1）に等しい表示時刻を持つオーディオサンプルを含んでいる。また、TS2のオーディオストリームの最初のオーディオフレームA_startは、TS2の最初の表示ピクチャの表示開始時（IN

_time2) に等しい表示時刻を持つオーディオサンプルを含んでいる。これにより、TS1とTS2との接続点において、オーディオプレゼンテーションユニットのシーケンスにギャップは存在せず、2オーディオフレーム区間未満のオーディオプレゼンテーションユニットの長さで定義されるオーディオオーバーラップ (Audio overlap) が生じる。接続点におけるTSは、後述するDVR-STD (Digital VideoRecording-System Target Decoder) に従ったDVR MPEG2TSである。

DVR-STDは、シームレス接続された2つのPlayItemによって参照されるAVストリームの生成及び検証の際におけるデコード処理をモデル化するための概念モデルである。DVR-STDモデル (DVR MPEG2 transport stream player model) を図8に示す。

図8に示すように、本実施の形態における情報処理装置 (DVR MPEG2 transport stream player model、以下プレーヤという。) 1は、シームレスに再生されるよう接続されたTSからトランスポートパケット (Transport packets) を読み出し出力する出力部10と、出力部10からのトランスポートパケットをデコードするデコーダ (DVR-STD) 20とから構成される。このデコーダ20は、後述するように、上述した従来のDVR-STDにおけるトランスポートパケットの入力タイミングとオーディオバッファの容量とを変更したものである。出力部10において、読出部 (DVRdrive) 11から読出しレート R_{UD} で読み出されたTSファイルは、リードバッファ12にバッファリングされ、このリードバッファ12からソースパケット (Source packets) がソースデパケットタイザ部 (source depacketizer) 13へビットレート R_{MAX} で読み出される。 R_{MAX} は、ソースパケットストリームのビットレートである。

パルス発振器 (27MHz X-tal) 14は、27MHzのパルスを発生する。アライバルクロックカウンタ (Arrival time clock counter) 15は、この27MHzの周波数のパルスをカウントするバイナリカウンタであり、ソースデパケットタイザ部13に、時刻 $t(i)$ におけるArrival time clock counterのカウント値 $Arrival_time_clock(i)$ を供給する。

上述したように、1つのソースパケットは、1つのトランスポートパケットとそのarrival_time_stampを持つ。現在のソースパケットのarrival_time_stamp

がarrival_time_clock(i)のLSB30ビットの値と等しいとき、ソースデパケッタイザ部13からそのトランスポートパケットが出力される。TS_recording_rateは、TSのビットレートである。

また、図8に示すn、TBn、MBn、EBn、TBsys、Bsys、Rxn、Rbxn、Rxsys、Dn、Dsys、On及びPn(k)の表記方法は、ISO/IEC13818-1 (MPEG2 Systems規格)のT-STD (ISO/IEC 13818-1で規定されるtransport stream system target decoder)に定義されているものと同じである。即ち、次の通りである。

n: エレメンタリストリームnのインデクス番号

TBn: エレメンタリストリームnのトランスポートバッファ

MBn: エレメンタリストリームnの多重バッファ (ビデオストリームについてのみ存在)

EBn: エレメンタリストリームnのエレメンタリストリームバッファ、ビデオストリームについてのみ存在

TBsys: 復号中のプログラムのシステム情報のための入力バッファ

Bsys: 復号中のプログラムのシステム情報のためのシステムターゲットデコーダ内のメインバッファ

Rxn: データがTBnから取り除かれる伝送レート

Rbxn: PESパケットペイロードがMBnから取り除かれる伝送レート (ビデオストリームについてのみ存在)

Rxsys: データがTBsysから取り除かれる伝送レート

Dn: エレメンタリストリームnのデコーダ

Dsys: 復号中のプログラムのシステム情報に関するデコーダ

On: ビデオストリームnの再配列バッファ (re-ordering buffer)

Pn(k): エレメンタリストリームnのk番目のプレゼンテーションユニット

次に、デコーダ20のデコーディングプロセスについて説明する。まず、1つのDVR MPEG2 TSを再生しているときのデコーディングプロセスを説明する。

単一のDVR MPEG2TSを再生している間は、トランスポートパケットをTB1、TBn又はTBsysのバッファへ入力するタイミングは、ソースパケットのarrival_time_stampにより決定される。

TB1、MB1、EB1、TBn、Bn、TBsys及びTBsysのバッファリング動作の規定は、ISO/IEC 13818-1に規定されているT-STDと同じである。復号動作と表示動作の規定もまた、ISO/IEC 13818-1に規定されているT-STDと同じである。

次に、シームレス接続されたPlayItemを再生している間のデコーディングプロセスについて説明する。図9は、あるAVストリーム(TS1)から、これにシームレスに接続された次のAVストリーム(TS2)へと移るときのトランスポートパケットの入力、復号、及び表示のタイミングチャートである。

ここでは、シームレス接続されたPlayItemによって参照される2つのAVストリームの再生について説明をすることにし、以後の説明では、図2又は図3に示した、シームレス接続されたTS1及びTS2の再生について説明する。したがって、TS1は、先行するストリームであり、TS2は、現在のストリームである。また、TS1及びTS2で区分された各パケットは、TS1及びTS2のソースパケットSP1、SP2を表している。

あるAVストリーム(TS1)からそれにシームレスに接続された次のAVストリーム(TS2)へと移る間には、TS2のアライバルタイムベースの時間軸(図9においてATC2で示される)は、TS1のアライバルタイムベースの時間軸(図9におけるATC1で示される)と同じ時間軸でない。また、TS2のシステムタイムベースの時間軸(図9においてSTC2で示される)は、TS1のシステムタイムベースの時間軸(図9においてSTC1で示される)と同じ時間軸でない。ビデオの表示は、シームレスに連続していることが要求される。オーディオのプレゼンテーションユニットの表示時間にはオーバーラップがあってもよい。

ここで、本実施の形態におけるプレーヤ1においては、上述した日本公開特許公報2000-175152号、日本公開特許公報2001-544118号及び日本公開特許公報2002-158974号に記載のプレーヤ101に対し、以下の2点を変更することで、オーディオバッファを最適な容量とするものである。まず、1つ目の変更点について説明する。1つ目の変更点は、あるAVストリーム(TS1)からそれにシームレスに接続された次のAVストリーム(TS2)へと移るとき、TS1の最後のパケットまでのデコーダ20への入力をそれらソースパケットのarrival_time_stampによって決定することとする点である。

即ち、上述した如く、従来は、TS1の最後のビデオパッケージがTB1に入力終了する時刻T1から、TS1の最後のバイトが入力終了するまでの時刻T2までの間は、arrival_time_stampを無視して、TSの最大ビットレートで、トランスポートパッケージがバッファに入力されていたのに対し、本実施の形態においては、T1からT2までの間のソースパッケージの入力を、時刻T1までと同様にして、TS1のソースパッケージのarrival_time_stampによって決定する。これにより、従来ソースパッケージのarrival_time_stampを無視して、TSの最大ビットレート R_{MAX} で入力するために必要とされていた1秒分の付加的なバッファは、不要となる。

この場合のデコーダ20への入力タイミングについて、図9を参照して説明する。

(1) 時刻T1までの時間

時刻T1までの時間、即ち、TS1の最後のビデオパッケージがデコーダ20のTB1に入力終了するまでは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパッケージSP1のarrival_time_stampによって決定される。

(2) 時刻T1からT2まで

TS1の残りのパッケージがデコーダ20へ入力するタイミングもまた、TS1のソースパッケージSP1のarrival_time_stampによって決定される。TS1の最後のバイトがバッファへ入力する時刻は、時刻T2である。

(3) 時刻T2以後

T2の時刻において、アライバルタイムクロックカウンタ15は、TS2の最初のソースパッケージのarrival_time_stampの値にリセットされる。デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS2のソースパッケージSP2のarrival_time_stampによって決定される。

即ち、デコーダ20への入力タイミングは、TS1の最後のバイトがバッファへ入力する時刻T2までは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングをTS1のソースパッケージSP1のarrival_time_stampによって決定し、T2以降は、TS2のソースパッケージSP2のarrival_time_stampによって決定する。

次に、ビデオのプレゼンテーションタイミングについて説明する。ビデオプレゼンテーションユニットの表示は、上述の図2又は図3に示すような接続点（コ

ネクションポイント)を通して、ギャップなしに連続しなければならない。即ち、TS1の最後のビデオデータ(第1のピクチャ)に、TS2の最初のビデオデータ(第2のピクチャ)が連続して再生される。ここで、

STC1: TS1のシステムタイムベースの時間軸(図9ではSTC1と示す。)

STC2: TS2のシステムタイムベースの時間軸(図9ではSTC2と示す。)正確には、STC2は、TS2の最初のPCR((Program Clock Reference))がT-STDに入力した時刻から開始される。

とする。

また、STC1とSTC2との間のオフセットは、次のように決定される。即ち、

$PTS1_{end}$: TS1の最後のビデオプレゼンテーションユニットに対応するSTC1上のPTS

$PTS2_{start}$: TS2の最初のビデオプレゼンテーションユニットに対応するSTC2上のPTS

T_{pp} : TS1の最後のビデオプレゼンテーションユニットの表示期間

とすると、2つのシステムタイムベースの間のオフセット値STC_deltaは、下記式(6)のように計算される。

$$STC_delta = PTS1_{end} + T_{pp} - PTS2_{start} \quad \dots (6)$$

次に、オーディオのプレゼンテーションタイミングについて説明する。TS1とTS2との接続点において、オーディオプレゼンテーションユニットの表示タイミングのオーバーラップがあってもよく、それは0以上であって2オーディオフレーム未満である(図9のオーディオオーバーラップを参照)。プレーヤには、どちらのオーディオサンプルを選択するかということ及び、オーディオプレゼンテーションユニットの表示を接続点の後の補正されたタイムベースに再同期する制御が要求される。

TS1からそれにシームレスに接続された次のTS2へと移るとき、デコーダ20のシステムタイムクロックの制御について、プレーヤが行う処理を説明する。

時刻 T 5 において、TS1の最後のオーディオプレゼンテーションユニットが表示される。システムタイムクロックは、時刻 T 2 から T 5 の間にオーバーラップしていてもよい。この区間では、デコーダ 2 0 は、システムタイムクロックを古いタイムベースの値 (STC1) を新しいタイムベースの値 (STC2) に切り替える。STC2の値は、下記式 (7) のように計算できる。

$$STC2 = STC1 - STC_delta \quad \dots (7)$$

次に、TS1から、このTS1にシームレスに接続される次のTS2へと移るとき、TS1及びTS2が満たさなければいけない符号化条件を説明する。ここで、
 $STC1^{1_end}$: TS1の最後のパケットの最後のバイトがデコーダ 2 0 へ到着するときのシステムタイムベースSTC1上のSTCの値
 $STC2^{2_start}$: TS2の最初のパケットの最初のバイトがデコーダ 2 0 へ到着するときのシステムタイムベースSTC2上のSTCの値
 $STC2^{1_end}$: $STC1^{1_end}$ の値をシステムタイムベースSTC2上の値に換算した値
 とすると、 $STC2^{1_end}$ は、下記式 (8) のようにして計算される。

$$STC2^{1_end} = STC1^{1_end} - STC_delta \quad \dots (8)$$

ここで、デコーダ 2 0 がDVR-STDに従うために、次の2つの条件を満たすことが要求される。

(条件 1)

TS2の最初のパケットのデコーダ 2 0 への到着タイミングは、次に示す不等式 (9) を満たさなければならない。

$$STC2^{2_start} > STC2^{1_end} \quad \dots (9)$$

この上記不等式 (9) を満たすように、Clip1及び／又はClip2の部分的なスト

リームを再エンコード及び／又は再多重化することが必要になる。

(条件 2)

STC1とSTC2とを同じ時間軸上に換算したシステムタイムベースの時間軸上において、TS1からのビデオパケットの入力とそれに続くTS2からのビデオパケットの入力は、ビデオバッファをオーバーフロー及びアンダフローさせてはならない。また、STC1とSTC2とを同じ時間軸上に換算したシステムタイムベースの時間軸上において、TS1からのパケットの入力とそれに続くTS2からのパケットの入力は、デコーダ20のすべてのバッファをオーバーフロー及びアンダフローさせてはならない。

図10は、あるAVストリーム(TS1)からそれにシームレスに接続された次のAVストリーム(TS2)へと移るときのトランスポートパケットの入力、復号、表示のタイミングチャートの他の例である。この場合もまた、TS1の最後のパケットまでのデコーダ20への入力をそれらソースパケットのarrival_time_stampによって決定することは同様であるが、図9に示すタイミングチャートとの一点の違いとして、図10に示すように、TS1の最後のパケットの直後にTS2の最初のパケットを入力する必要がないように、所定の時間間隔(δt_1 :時刻 $T_2 \sim T_2'$ の間)を設けている。これにより、TS2の最初のパケットの入力タイミングの決定が図9の場合よりも柔軟であるので、TS2の符号化を容易にする効果がある。

この場合のデコーダ20への入力タイミングについて、図10を参照して説明する。

(1) 時刻 T_2 までの時間

時刻 T_2 までの時間、即ち、TS1の最後のパケットの最後のバイトがデコーダ20へ入力終了するまでは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパケットSP1のarrival_time_stampによって決定される。

(2) 時刻 T_2' 以後

時刻 T_2 から δt_1 の時間の後、時刻 T_2' の時刻において、アライバルタイムクロックカウンタ15は、TS2の最初のソースパケットのarrival_time_stampの値にリセットされる。デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミ

ングは、TS2のソースパケットSP2のarrival_time_stampによって決定される。

ここで、図10に示すように、delta₁を設ける場合、上述のSTC2²_{start}とSTC2¹_{end}は下記の関係式(10)を満たさなければならない。

$$STC2^{2}_{start} > STC2^{1}_{end} + \delta_{1} \quad \dots (10)$$

図11は、あるAVストリーム(TS1)からそれにシームレスに接続された次のAVストリーム(TS2)へと移るときのトランスポートパケットの入力、復号、表示のタイミングチャートの他の例である。この場合もまた、TS1の最後のパケットまでのデコーダ20への入力をそれらソースパケットのarrival_time_stampによって決定することは同様であるが、図10に示すタイミングチャートとの一点の違いとして、図11に示すように、所定の時間間隔(ATC_delta:時刻T2~T2'の間)を設けている。これにより、TS2の最初のパケットの入力タイミングの決定が図9の場合よりも柔軟であるので、TS2の符号化を容易にする効果がある。

この場合のデコーダ20への入力タイミングについて、図11を参照して説明する。

(1) 時刻T2までの時間

時刻T2までの時間、即ち、TS1の最後のパケットの最後のバイトがデコーダ20へ入力終了するまでは、デコーダ20のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS1のソースパケットSP1のarrival_time_stampによって決定される。

(2) 時刻T2からT2'までの時間

時刻T2'は、TS2の最初のパケットがデコーダ20に入力される時刻である。ATC_deltaは、TS1の最後のパケットのarrival_time_stamp時刻(ATC1上の時刻)からATC1上に投射された時刻T2'までのオフセット時間である。

(3) 時刻T2'以後

時刻T2'の時刻において、アライバルタイムクロックカウンタ15は、TS2の最初のソースパケットのarrival_time_stampの値にリセットされる。デコーダ2

0のTB1、TBn又はTBsysのバッファへの入力タイミングは、TS2のソースパケットSP2のarrival_time_stampによって決定される。

ここで、ATC_deltaの値は、上記式(6)のSTC_deltaを満たすように値が決定される。

このATC_deltaの値は、ストリームデータの付属情報として管理される。TS1とTS2とが図11のようにシームレスに接続される場合、ATC_deltaの値は、TS1の付属情報として管理される。

図12には、ATC_deltaを格納するための付属情報ClipInfo()のデータフォーマットを示す。

図12において、is_ATC_deltaは、ClipInfo()がATC_deltaの値を持つか否かを示すフラグである。ClipInfo()には複数の値を登録することができる。これは、TS1に接続されるところのTS2を図13に示すように複数持たせることができるようにするためである。is_ATC_deltaフラグが1である場合、number_of_ATC_delta_entriesは、このClipInfo()に登録されているATC_deltaの個数を示す。

また、図12において、following_Clip_Information_file_nameは、TS1に接続されるところのTS2のストリームの名前である。following_Clip_Information_file_nameに対応するところのTS2が複数存在する場合、それぞれのTS2に対するATC_deltaの値がClipInfo()に登録される。

図8のDVR-STDモデルにTS1及びTS2が入力される場合、それぞれの多重化ストリームとともに、その付属情報ClipInfo()が入力される。ClipInfo()は、上述のATC_deltaの情報を含み、ATC_deltaは、DVR-STDモデルのコントローラ(図8で図示せず)によって、上述のようにTS1とTS2との切替わりにおいて所定の方法で取り扱われる。

2つ目の変更点としては、デコーダ20のオーディオバッファのサイズを、次の条件が満たされるような十分な大きさに変更する。この条件とは、TS1からこれにシームレスに接続される次のTS2へと移るとき、TS1の最後のトランスポートパケットを入力後に、TS2の最初にデコードされるピクチャ(Iピクチャ)をそのデコードタイミングまでにビデオバッファへ入力できることである。

この条件を満たすために、オーディオバッファの必要な容量の最大値は次のよ

うな値である。即ち、「Iピクチャの最大のビット量をそのデコードタイミングまでにビデオバッファへ入力できる時間」に相当する長さのオーディオのデータ量を蓄えられる大きさである。オーディオバッファの必要量の最大値 EBn_max は、次の式(11)で計算できる。

$$EBn_max = (I_max / Rv) * Ra \text{ [bits]} \quad \dots (11)$$

ここで、 I_max は、Iピクチャの最大のビット量であり、これは、図8に示すビデオコードバッファEB1の大きさである。 Rv は、ビデオコードバッファEB1への入力ビットレートである。また、 Ra は、オーディオストリームのビットレートである。上記式(11)に示すように、求めるべきオーディオバッファの大きさ EBn_max は、ビデオエレメンタリストリームバッファ(EB1)への入力ビットレートでビデオコードバッファEB1のバッファ占有量をゼロから I_max にするまでにかかる時間(I_max / Rv)に Ra を掛けた値である。

また、具体的な値として、少なくとも100ミリ秒分のオーディオデータを蓄えられるバッファサイズを推奨する。これは以下の理由による。即ち、Iピクチャを0.5秒ごとに符号化する場合、Iピクチャのビットサイズは、一般的に、符号化ビットレートの10%以下である。例えば、符号化ビットレートが10Mbpsの場合、Iピクチャのサイズは、通常1Mbits以下である。

したがって、第1に、少なくとも、100ミリ秒分の時間があれば、Iピクチャをそのデコードタイミングまでにデコーダ20のビデオバッファへ入力できる。また、第2にデコーダ20のオーディオバッファが、100ミリ秒分のオーディオデータを蓄えられるのであれば、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く、オーディオバッファに入力終了するようにTS1を多重化できる。したがって、オーディオバッファを、少なくとも100ミリ秒分のオーディオデータを蓄えられるバッファサイズとすれば、上記第1及び第2の理由により、TS1からそれにシームレスに接続される次のTS2へと移るとき、TS1の最後のトランスポートパケットを入力後に、TS2の最初にデコードされるピク

チャ（Iピクチャ）をビデオバッファへ入力終了する（Iピクチャのコードタイミング）までの時間として少なくとも100ミリ秒分を確保することができる。

100ミリ秒分のオーディオデータを蓄えられるオーディオバッファの容量を具体的に計算すると次のようになる。

640 kbpsのドルビーAC3オーディオストリームの場合： $640 \text{ kbps} \times 0.1 \text{ s}$
 $\text{ec} = 64 \text{ kbitS} = 8 \text{ kBytes}$

Linear PCMオーディオストリーム、24bitSample、96 kHz sampling frequency、8 channelsの場合： $(24 \text{ bitSample} \times 96000 \text{ samples/sec} \times 8 \text{ ch}) \times 0.1 \text{ sec} = 230400 \text{ Bytes}$

次に、以上説明した本実施の形態のデコーダ20のように、DVR-STDのオーディオバッファのサイズを、「TS2の最初にデコードされるピクチャ（Iピクチャ）をそのデコードタイミングまでにビデオバッファへ入力できる時間に相当するオーディオのデータ量を蓄えられる大きさ」に変更することによる効果を図14及び図15を用いて更に詳細に説明する。

例として、オーディオストリームが、ビットレート640 kbps、サンプリング周波数48 kHzのAC3オーディオストリームの場合を説明する。AC3オーディオストリームの1オーディオフレームのサンプル数は1536 samplesであるので、その時間長は、32ミリ秒である。また、1オーディオフレームのバイトサイズは、2560 Bytesである。

図14の（a）及び（b）は、従来のDVR-STDの場合であり、オーディオのバッファサイズが4 kBytesである場合に、TS1からそれにシームレスに接続される次のTS2へと移るときのDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。この図14において、TS1のビデオ／オーディオデータのバッファ遷移を破線で示し、また、TS2のビデオ／オーディオデータのバッファ遷移を実線で示す。

4 kBytesのオーディオバッファは、50ミリ秒分のオーディオデータを蓄えられる。したがって、TS1の最後のオーディオパケットの最後のバイトがDVR-STDに到着する時刻であるSTC1¹_{audio_end}において、オーディオデータをそれが再生されるタイミングよりも50ミリ秒だけ早く入力終了するようにTS1を多重化できる。

しかしながら、50ミリ秒は、TS2の最初にデコードされるピクチャ（Iピクチャ）をそのデコードタイミングまでにビデオバッファへ入力するためには、不十分である。そのため、TS2の最初にデコードされるピクチャ（Iピクチャ）のサイズを小さくするようにエンコードを制限することになり、画質が悪くなる問題がある。

即ち、4 kBytesのオーディオバッファでは、先送りすることができるオーディオデータは50ミリ秒であるので、TS2の最初のIピクチャをビデオバッファに入力するための時間である図14に示すスタートアップディレイ（start-up delay） t_1 は、最大50ミリ秒と小さくなる。したがって、TS2の最初のIピクチャを入力する時間を十分とれず、Iピクチャサイズ S_1 が小さくなり、符号化が制限されIピクチャの画質が劣化する。なお、このstart-up delayを大きくするため、上述した如く、従来は、4 kBytesに加えて1秒分程度の付加的なバッファを設け、且つ $T_1 \sim T_2$ の間はTSの最大レート RM_{AX} で入力する必要があった。ここでは、ビットレート640 kbpsのAC3オーディオストリームについて説明しているが、上述した如く、マルチチャンネルのLPCMオーディオに対応可能なように設計すると、1秒分程度の付加的なバッファは極めて大きくなってしまう。

この問題を解決するために、本実施の形態におけるデコーダ20のように、DVR-STDのオーディオバッファサイズを例えば8 kBytesに変更する。図15の（a）及び（b）は、オーディオバッファの容量を最適化した例を示すものであって、オーディオのバッファサイズが8 kBytesである場合に、TS1からこれにシームレスに接続される次のTS2へと移るときの本実施の形態におけるDVR-STDの夫々ビデオバッファ及びオーディオバッファのビット占有量の変化の例を示すグラフ図である。この図15において、TS1のビデオ／オーディオデータのバッファ遷移を破線で示し、また、TS2のビデオ／オーディオデータのバッファ遷移を実線で示す。

8 kBytesのオーディオバッファは、100ミリ秒分のオーディオデータを蓄えられる。したがって、TS1の最後のオーディオパケットの最後のバイトがDVR-STDに到着する時刻である $STC1^{audio_end}$ において、オーディオデータをそれが再生されるタイミングよりも100ミリ秒だけ早く入力終了するようにTS1を多重化で

きる。少なくとも、100ミリ秒あれば、TS2の最初にデコードされるピクチャ（Iピクチャ）をそのデコードタイミングまでにビデオバッファへ入力する余裕ができる。即ち、TS2の最初のIピクチャを入力する時間（スタートアップディレイ） t_2 を十分とることができ、そのため、TS2の最初にデコードされるピクチャ（Iピクチャ）サイズ S_2 を大きくすることができ、したがってIピクチャのエンコード制限が小さく、高画質にすることができる。

また、図8に示すようなプレーヤモデル1において、トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、そのアライバルタイムスタンプに基づきデコーダにより読み出されてデコードされるTSは、多重化装置（情報処理装置）において生成され記録されたものとして行うことができる。

多重化装置は、例えば上述の図4乃至6を参照して説明したように、所定のピクチャで表示終了するよう、再符号化したClip1（第1のビデオ符号化ストリーム）を生成し、このピクチャに続けて表示され、且つ表示開始できるよう再符号化したClip2（第2のビデオ符号化ストリーム）を生成するビデオ符号化部と、Clip1とこのClip1に同期したオーディオ符号化ストリームとを多重化してTS1を生成し、Clip2とこのClip2に同期したオーディオ符号化ストリームとを多重化してTS2を生成する多重化部と、TS1及びTS2からなる多重化ストリームを記録する記録部とを備える。ここで、多重化部においては、上記第2のピクチャであるIピクチャをデコーダ20のビデオバッファへ入力するために要する時間分のオーディオデータを当該Iピクチャがデコーダ20へ入力開始される前までに入力終了可能なように多重化する。なお、図5に示すように、符号化部においてBridge-Clipを生成し、多重化部においてBridge-Clipも合わせて多重化するようにしてもよいことはもちろんである。

このような多重化装置により生成された多重化ストリームが記録された記録媒体には、第1のピクチャで表示終了するTS1と、この第1のピクチャに続けて再生する第2のピクチャから表示開始するTS2とから生成され、TS1及びTS2が夫々のアライバルタイムスタンプに基づいてデコーダ20へ入力可能であって、且つ、第2のピクチャであるTS2の最初のピクチャをデコーダへ入力するために要する時間

分のオーディオデータを第2のピクチャがデコーダ20に入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録されたものとなっている。

このように構成された本実施の形態においては、シームレスに接続されたTS1及びTS2を再生する際、TS1の最後のビデオパッケージがデコーダ20のTB1に入力終了した後からTS1の残りのパッケージがデコーダ20へ入力するまでにおいても、トランスポートパッケージをアライバルタイムスタンプに従って入力するようにし、オーディオバッファのサイズを、従来のDVR-STDにおける4k Bytesから、Iピクチャの最大のビット量をそのデコードタイミングまでにビデオバッファへ入力できる時間に相当する長さのオーディオのデータ量を蓄えられる大きさに変更することにより、TS1の最後のパッケージが入力終了してからTS2の最初のピクチャであるIピクチャをそのデコードタイミングまでに入力する時間（スタートアップディレイ）を十分確保することができるので、Iピクチャの符号化制限を小さくし、高画質とすることができる。

また、従来のように付加的なバッファを設ける方法であると、例えばTSにおけるオーディオデータをマルチチャンネルのLPCMオーディオデータとした場合に極めて大きい容量の付加的なバッファが必要になるのに対し、本実施の形態においては、オーディオバッファの容量を上記のように変更し、アライバルタイムスタンプに従ってトランスポートパッケージを入力するようにすることで、従来必要となっていた付加的なバッファを不要とすることができる。

なお、本発明は、図面を参照して説明した上述の実施例に限定されるものではなく、添付の請求の範囲及びその主旨を逸脱することなく、様々な変更、置換又はその同等のものを行うことができることは当業者にとって明らかである。

産業上の利用可能性

上述した本発明によれば、ビデオストリームとオーディオストリームの多重化ストリームをビデオフレーム精度で編集して、編集点をシームレスに再生することが可能であり、従来ソースパッケージのarrival_time_stampを無視して、TSの最

大ビットレートで入力するために必要とされていた1秒分の付加的なバッファを不要とし、従来よりもデコーダに必要なバッファ量を小さくできると共に、オーディオバッファのサイズを、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量とすることにより第2のピクチャのエンコード制限が小さく、高画質とすることができる。

請求の範囲

1. トラnsポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理装置において、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力手段と、

上記ソースパケットのうちビデオデータをバッファリングするビデオバッファと、

上記ソースパケットのうちオーディオデータをバッファリングするオーディオバッファと、

上記ビデオバッファにバッファリングされたビデオデータを復号するビデオ復号手段と、

上記オーディオバッファにバッファリングされたオーディオデータを復号するオーディオ復号手段とを有し、

上記オーディオバッファは、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータをバッファリング可能な容量を有する情報処理装置。

2. 上記オーディオバッファに必要な容量を EBn_max (bits) とし、上記第2のピクチャのビット量を I_max (bits) とし、上記ビデオバッファへの入力ビットレートを Rv (bps) とし、オーディオデータのビットレートを Ra (bps) としたとき、

$$EBn_max = (I_max / Rv) \times Ra$$

を満たす請求の範囲第1項記載の情報処理装置。

3. 上記第2のピクチャは、フレーム内符号化画像である請求の範囲第1項記載の情報処理装置。

4. 上記オーディオバッファは、少なくとも100ミリ秒分のオーディオデータをバッファリング可能な容量を有する請求の範囲第1項記載の情報処理装置。

5. 上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値STC1¹_{end}を上記時間差STC_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値をSTC2¹_{end} (=STC1¹_{end} - STC_delta) とし、上記第2の多重化ストリームの最初のソースパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値をSTC2²_{start}としたとき、上記多重化ストリームは、

$$STC2^{2}_{start} > STC2^{1}_{end}$$

を満たす請求の範囲第1項記載の情報処理装置。

6. 上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC_deltaとし、上記第1の多重化ストリームの最後のソースパケットの最後のバイトが上記出力手段から出力される該第1の多重化ストリームの時間軸上の値STC1¹_{end}を上記時間差STC_deltaにより上記第2の多重化ストリームの時間軸上の値に換算した値をSTC2¹_{end} (=STC1¹_{end} - STC_delta) とし、上記第2の多重化ストリームの最初のソースパケットの最初のバイトが上記出力手段から出力される該第2の多重化ストリームの時間軸上の値をSTC2²_{start}とし、上記第1の多重化ストリームの最後のソースパケットが上記出力手段から出力された後、所定時間delta1の経過後に上記第2の多重化ストリームの最初のソースパケットを上記出力手段から出力するものとしたとき、上記多重化ストリームは、

$$STC2^{2}_{start} > STC2^{1}_{end} + \text{delta1}$$

を満たす請求の範囲第1項記載の情報処理装置。

7. 上記第1の多重化ストリームの時間軸における上記第1のピクチャの表示終了時刻と上記第2の多重化ストリームの時間軸における上記第2のピクチャの表示開始時刻との時間差をSTC_deltaとし、上記第1の多重化ストリームの最後のソースパケットの出力を開始した後、所定時間ATC_deltaの経過後に上記第2の多重化ストリームの最初のソースパケットを上記出力手段から出力するものとし、

上記所定時間ATC_deltaは、上記時間差STC_deltaの値を満たすように決定されたものであり、

上記多重化ストリームは、上記時間差STC_deltaの値を満たすように多重化されたものである請求の範囲第1項記載の情報処理装置。

8. 上記所定時間ATC_deltaの値を上記第1の多重化ストリームの付属情報として管理する範囲第7項記載の情報処理装置。

9. 上記第1及び第2の多重化ストリームにおけるオーディオデータは、マルチチャンネルのオーディオデータからなる請求の範囲第1項記載の情報処理装置。

10. トラnsポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する情報処理方法において、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、

上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングする情報処理方法。

11. 上記オーディオバッファの必要ビット量をEBn_max (bits) とし、上記第2のピクチャのビット量をI_max (bits) とし、上記ビデオバッファへの入力ビットレートをRv (bps)、オーディオデータのビットレートをRa (bps) としたとき、

$$EBn_max = (I_max / Rv) \times Ra$$

を満たす請求の範囲第10項記載の情報処理方法。

1 2. トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する処理をコンピュータに実行させるためのプログラムであって、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、

上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングするプログラム。

1 3. 上記オーディオバッファの必要ビット量を EBn_max (bits) とし、上記第2のピクチャのビット量を I_max (bits) とし、上記ビデオバッファへの入力ビットレートを Rv (bps)、オーディオデータのビットレートを Ra (bps) としたとき、

$$EBn_max = (I_max / Rv) \times Ra$$

を満たす請求の範囲第1 2項記載のプログラム。

1 4. トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続された多重化ストリームを復号する処理をコンピュータに実行させるためのプログラムが記録されたコンピュータ読取可能な記録媒体であって、

上記多重化ストリームの上記アライバルタイムスタンプに従って上記ソースパケットを出力する出力工程と、

上記ソースパケットのうちビデオデータをビデオバッファにバッファリングし、オーディオデータをオーディオバッファにバッファリングするバッファリング工程と、

上記ビデオバッファ及びオーディオバッファにバッファリングされた夫々ビデオデータ及びオーディオデータを復号する復号工程とを有し、

上記バッファリング工程では、上記第2のピクチャを上記ビデオバッファに入力するために要する時間分のオーディオデータを、上記第2のピクチャを上記ビデオバッファにバッファリングする前に上記オーディオバッファにバッファリングするプログラムが記録された記録媒体。

15. 上記オーディオバッファの必要ビット量を EBn_max (bits) とし、上記第2のピクチャのビット量を I_max (bits) とし、上記ビデオバッファへの入力ビットレートを Rv (bps)、オーディオデータのビットレートを Ra (bps) としたとき、

$$EBn_max = (I_max / Rv) \times Ra$$

を満たす請求の範囲第14項記載の記録媒体。

16. トランスポートパケットとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなる多重化ストリームが記録された記録媒体であって、

上記多重化ストリームは、第1の多重化ストリームの最後のピクチャである第1のピクチャに第2の多重化ストリームの最初のピクチャである第2のピクチャが連続して再生されるよう接続され、上記第1及び第2の多重化ストリームが夫々のアライバルタイムスタンプに基づいてデコーダに入力可能であって、且つ、上記第2のピクチャを該デコーダに入力するために要する時間分のオーディオデータを該第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化された多重化ストリームが記録された記録媒体。

17. 上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータは、上記第2のピクチャのビット量を I_max (bits) とし、上記デコーダのビデオバッファへの入力ビットレートを Rv (bps)、オーディオデータのビットレートを Ra (bps) としたとき、 $(I_max / Rv) \times Ra$ である請求の範囲第16項記載の記録媒体。

18. 上記第2のピクチャは、フレーム内符号化画像である請求の範囲第16項記載の記録媒体。

19. транспорт пакетとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づきデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理装置において、

第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化手段と、

上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化手段とを有し、

上記多重化手段は、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化する情報処理装置。

20. 上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータは、上記第2のピクチャのビット量を I_{\max} (bits) とし、上記デコーダのビデオバッファへの入力ビットレートを R_v (bps)、オーディオデータのビットレートを R_a (bps) としたとき、 $(I_{\max}/R_v) \times R_a$ である請求の範囲第19項記載の情報処理装置。

21. 上記第2のピクチャは、フレーム内符号化画像である請求の範囲第19項記載の情報処理装置。

22. транспорт пакетとそのアライバルタイムスタンプとを有するソースパケットを単位とするデータ列からなり、該アライバルタイムスタンプに基づ

きデコーダにより読み出されてデコードされる多重化ストリームを生成する情報処理方法において、

第1のピクチャで表示終了する第1のビデオ符号化ストリームを生成し、この第1のピクチャに続けて表示される第2のピクチャから表示開始する第2のビデオ符号化ストリームを生成するビデオ符号化工程と、

上記第1のビデオ符号化ストリームとこの上記第1のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第1の多重化ストリームを生成し、上記第2のビデオ符号化ストリームとこの第2のビデオ符号化ストリームに同期したオーディオ符号化ストリームとを多重化して第2の多重化ストリームを生成し、上記第1の多重化ストリームの最後のピクチャである上記第1のピクチャに上記第2の多重化ストリームの最初のピクチャである上記第2のピクチャが連続して再生されるよう接続された多重化ストリームを生成する多重化工程とを有し、

上記多重化工程では、上記第2のピクチャを上記デコーダに入力するために要する時間分のオーディオデータを上記第2のピクチャが該デコーダに入力開始される前までに入力終了可能なように多重化する情報処理方法。

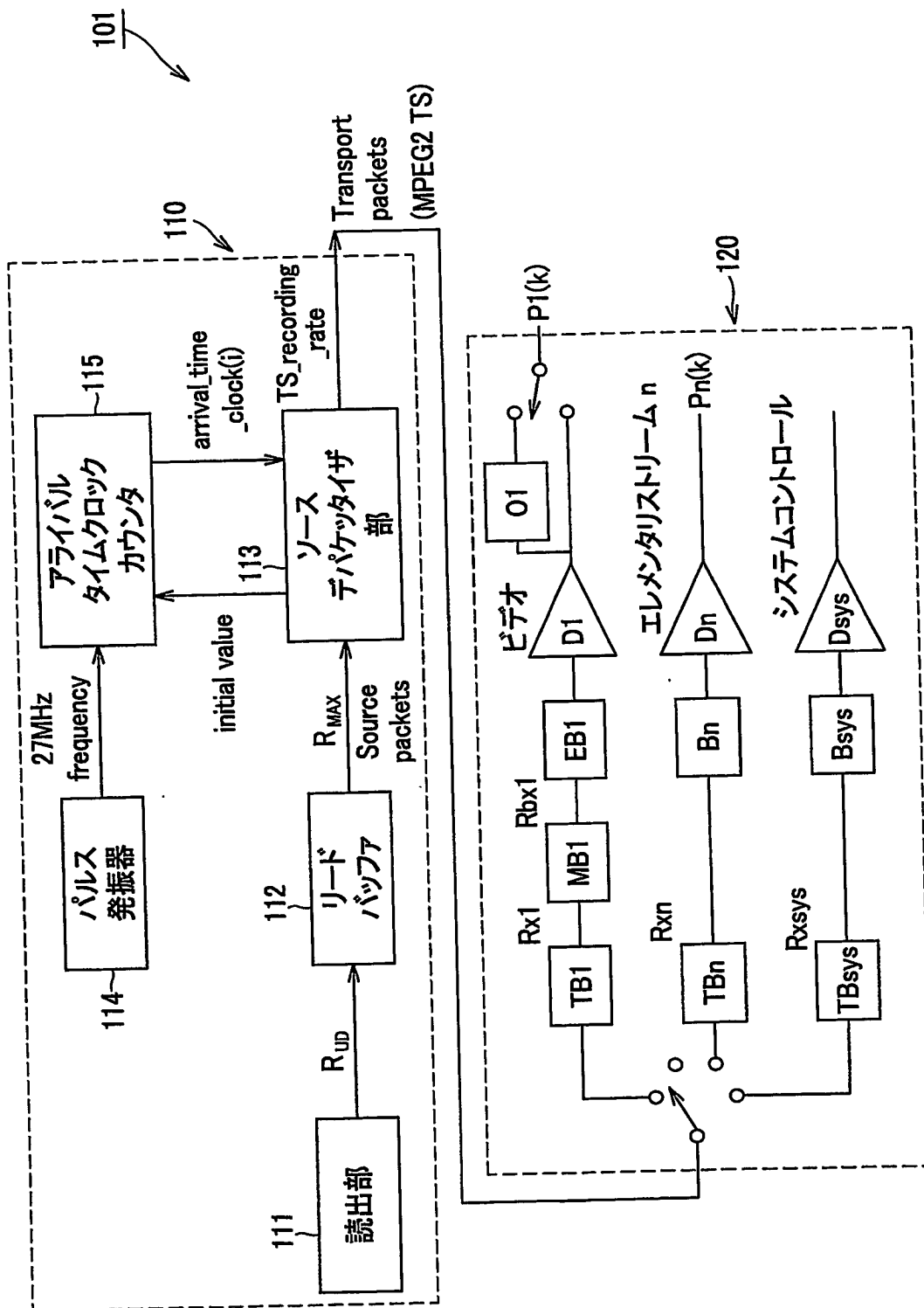


FIG.1

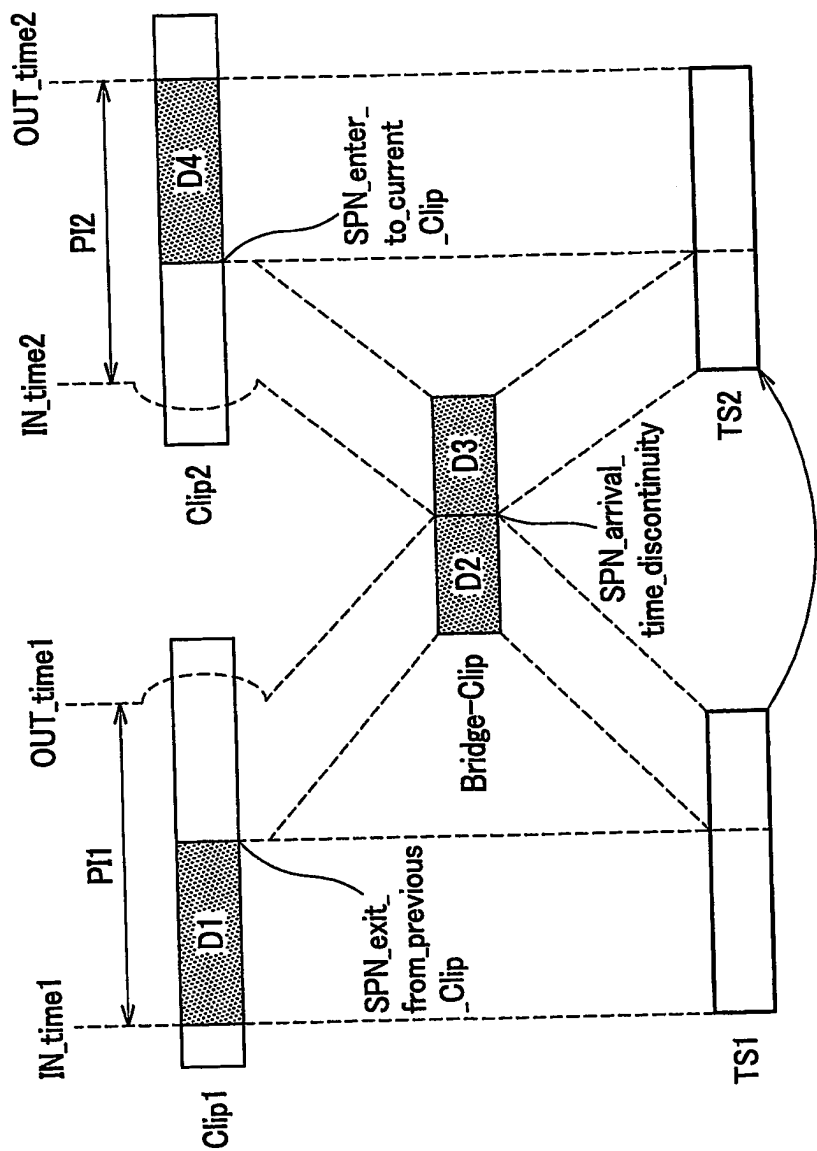


FIG.2

3/15

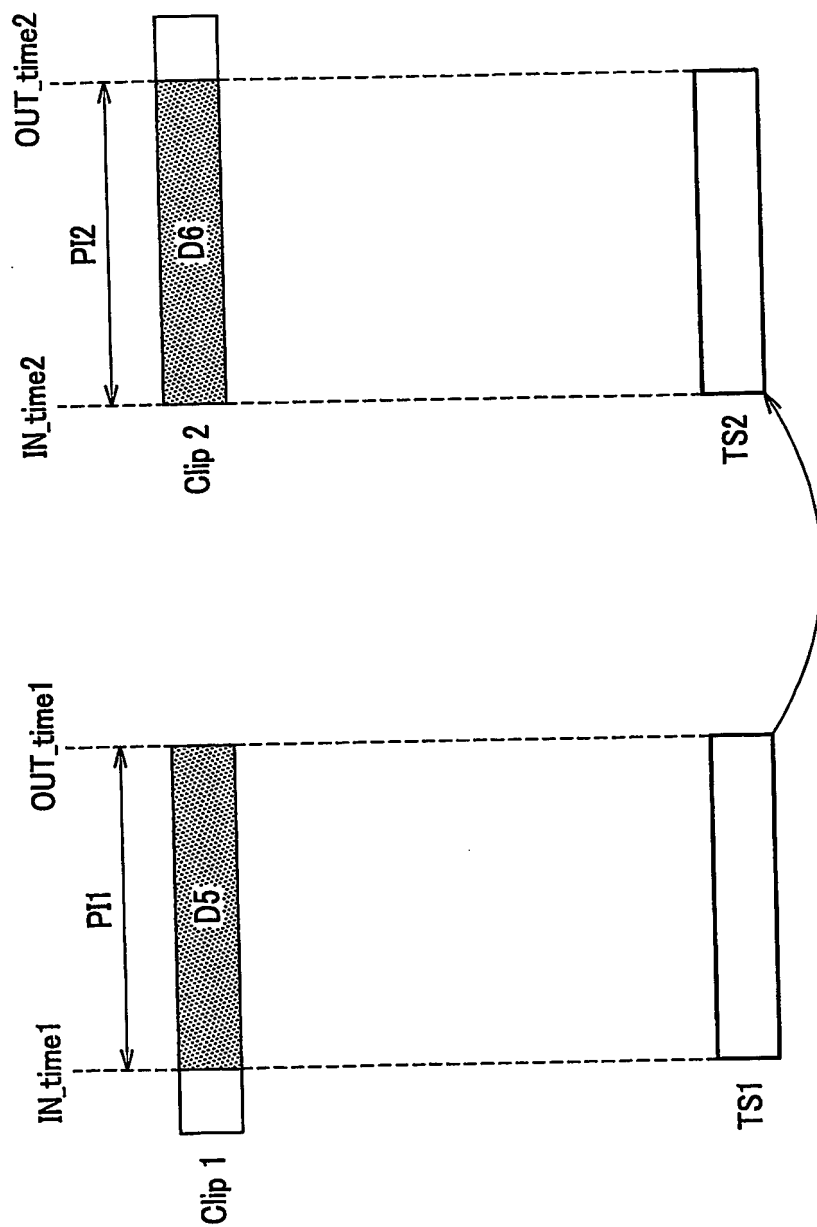


FIG.3

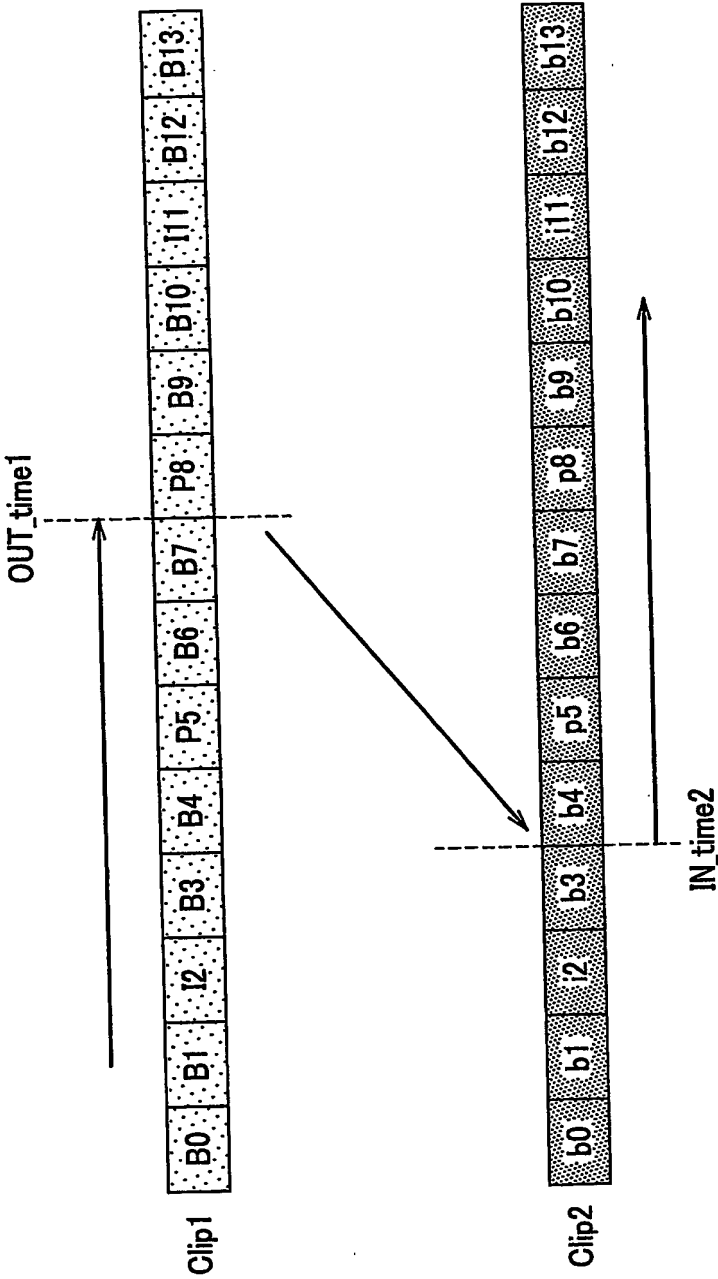


FIG.4

5/15

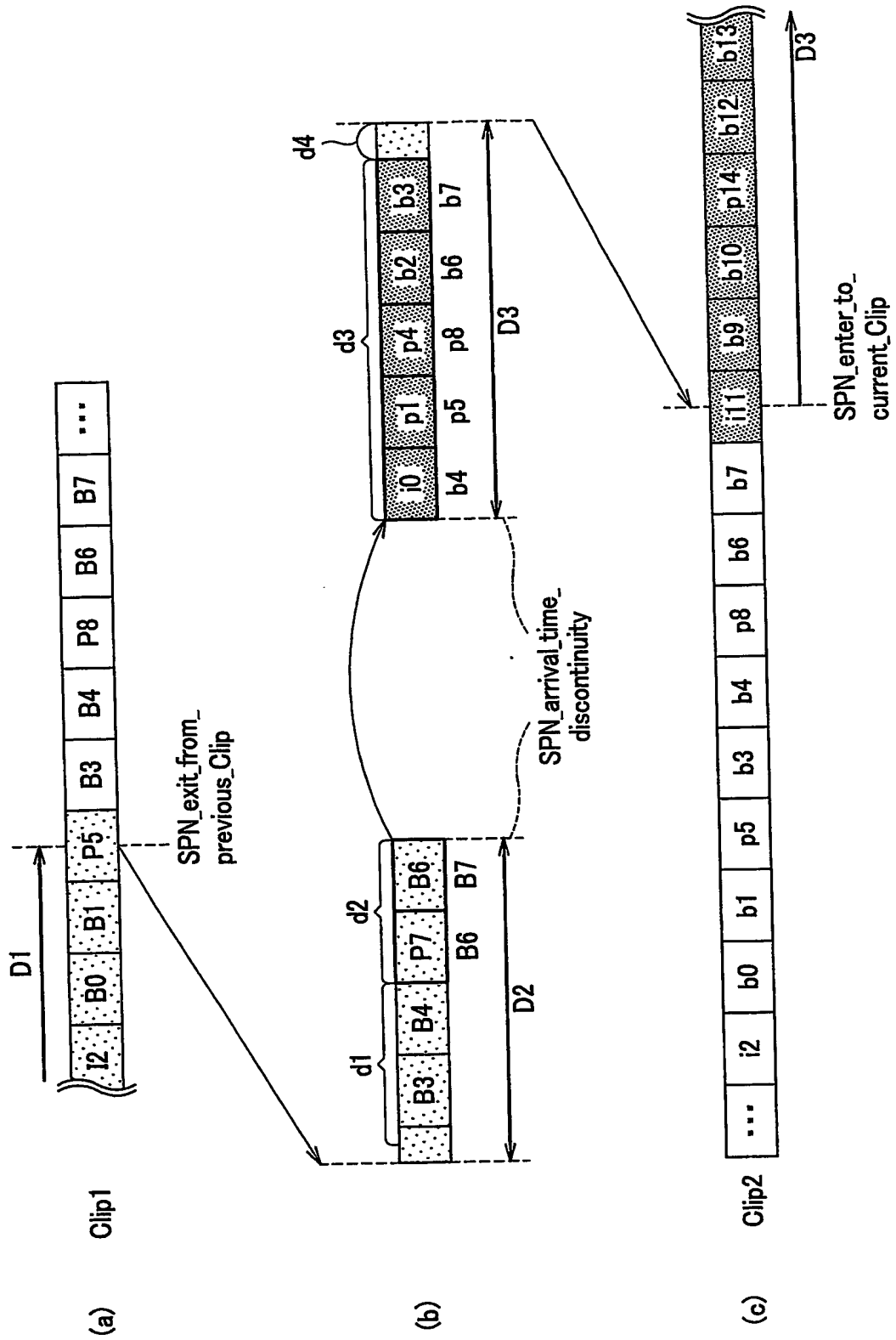


FIG.5

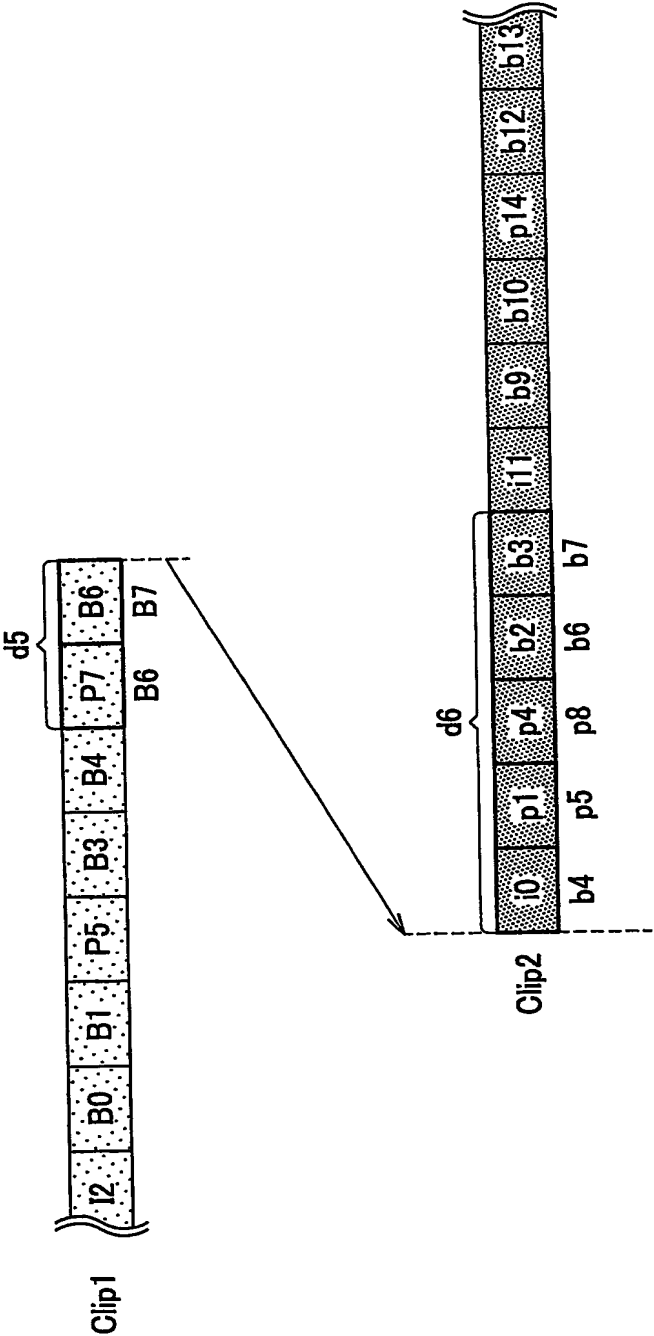


FIG.6

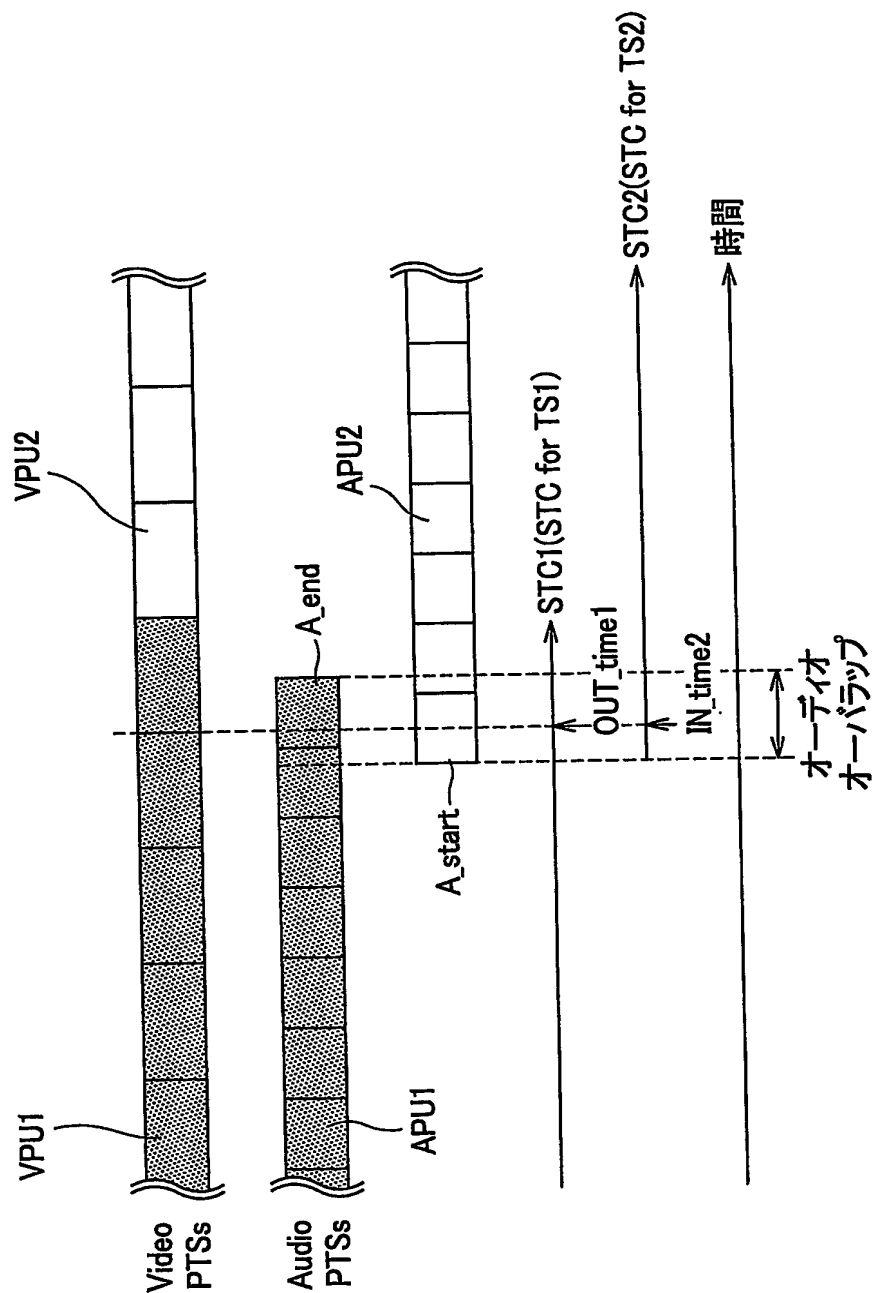


FIG.7

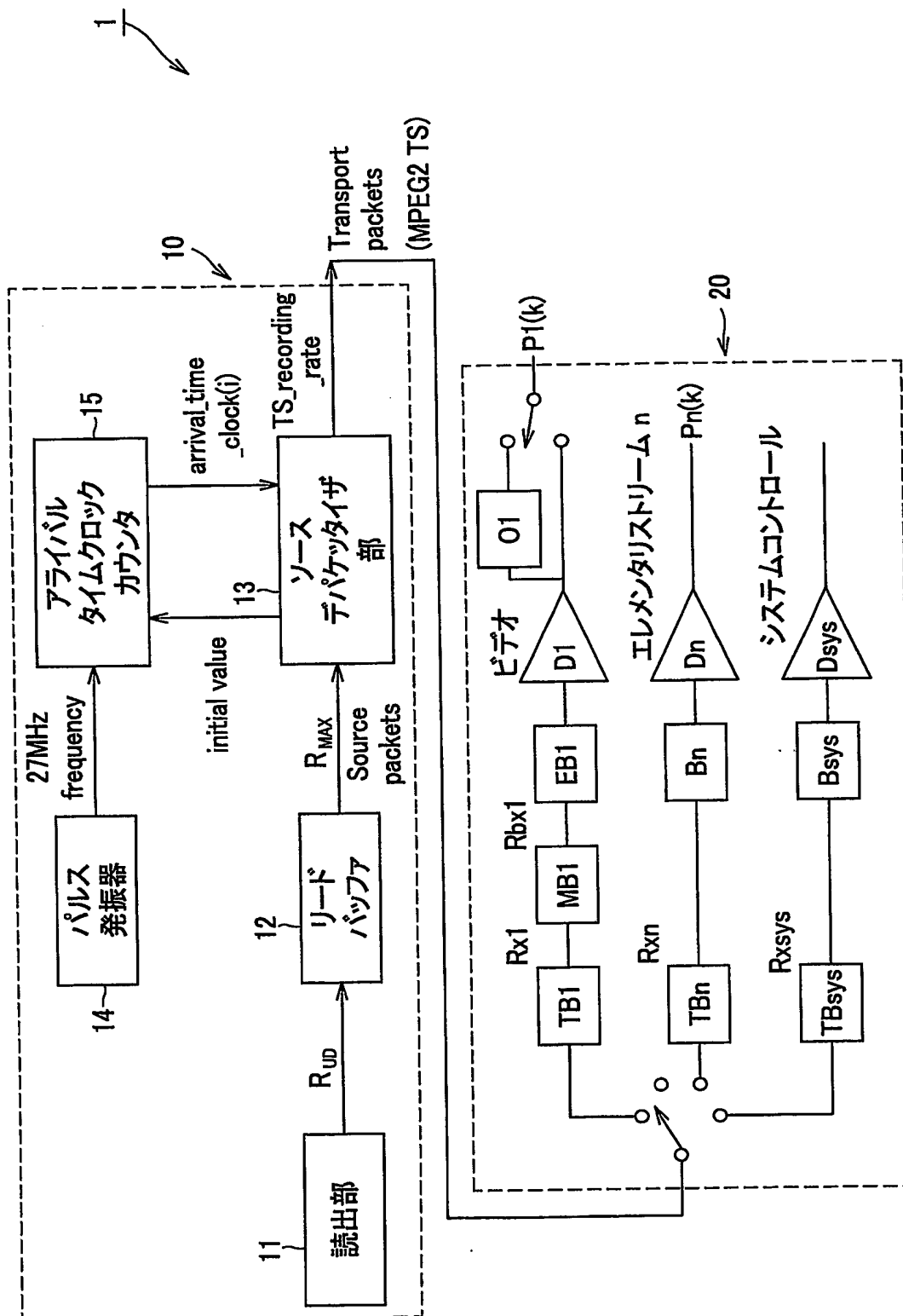


FIG.8

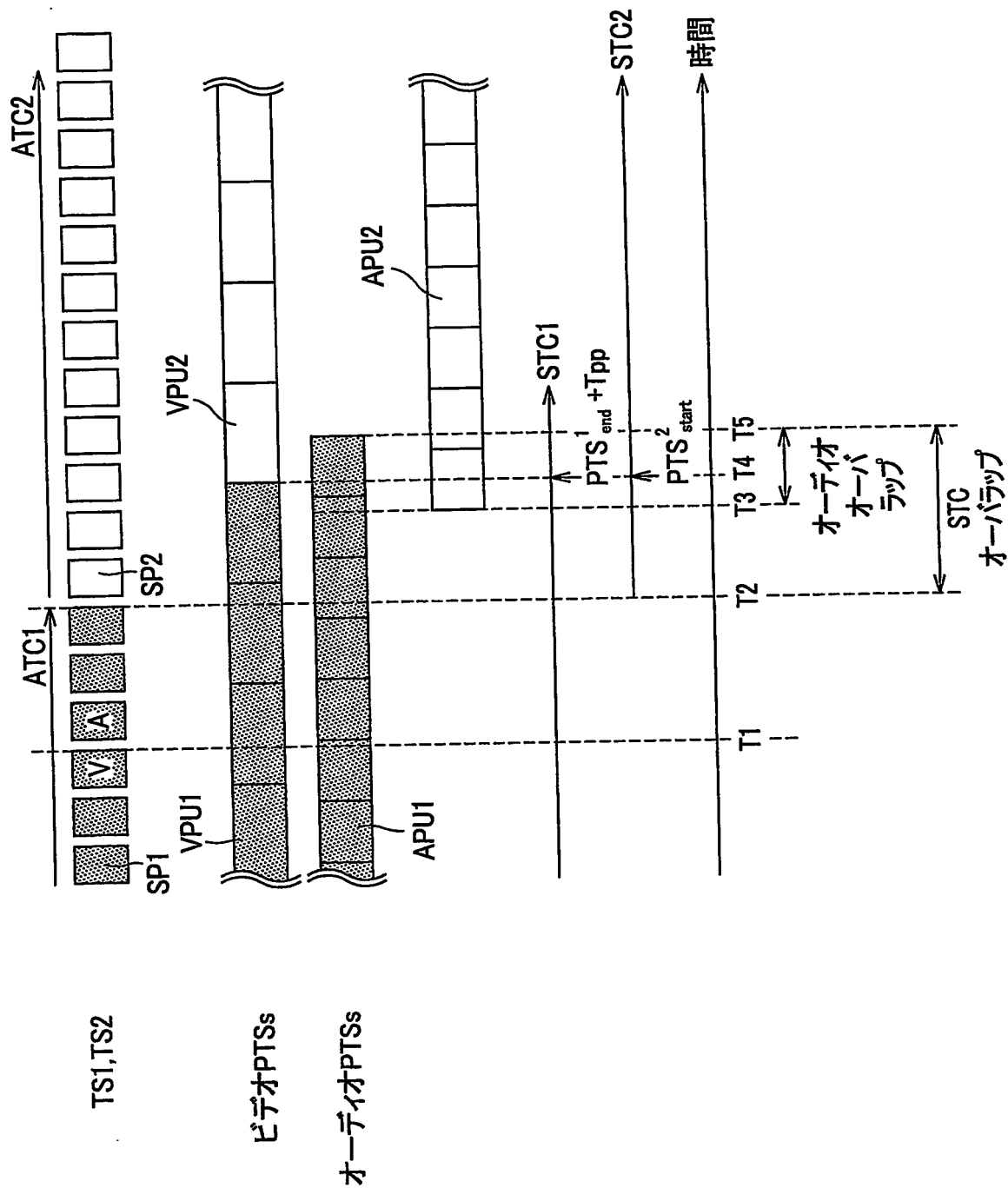


FIG.9

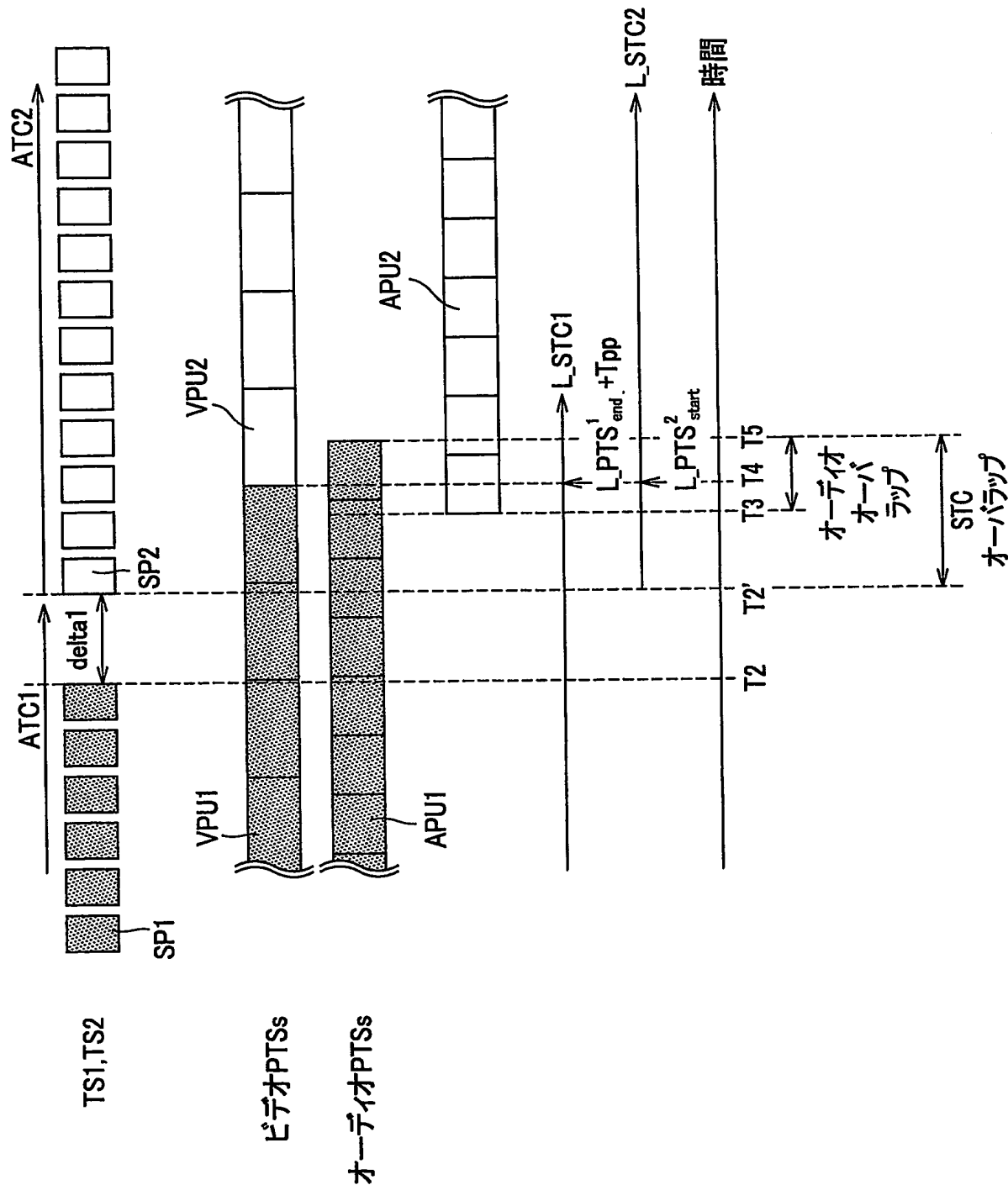


FIG.10

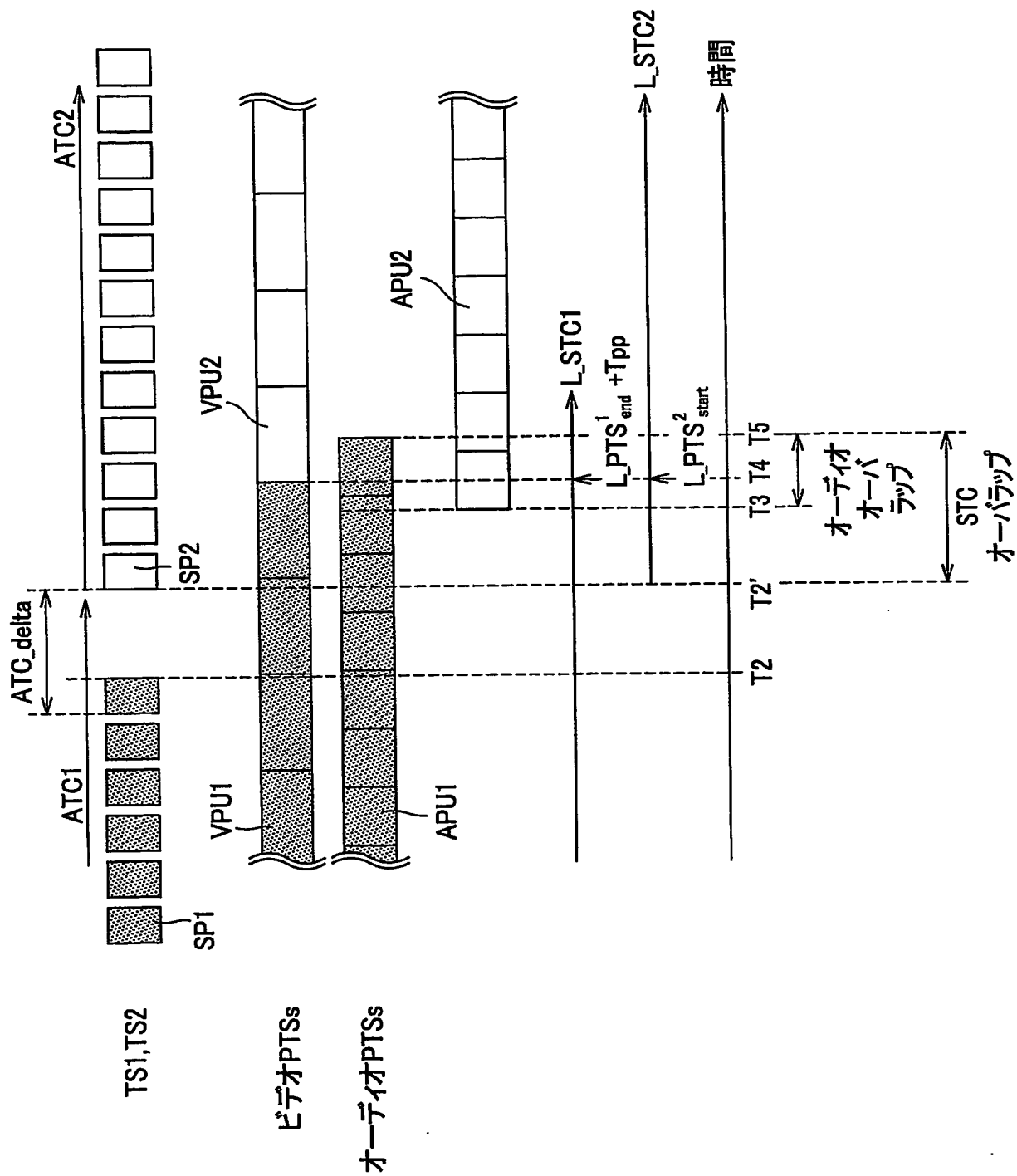
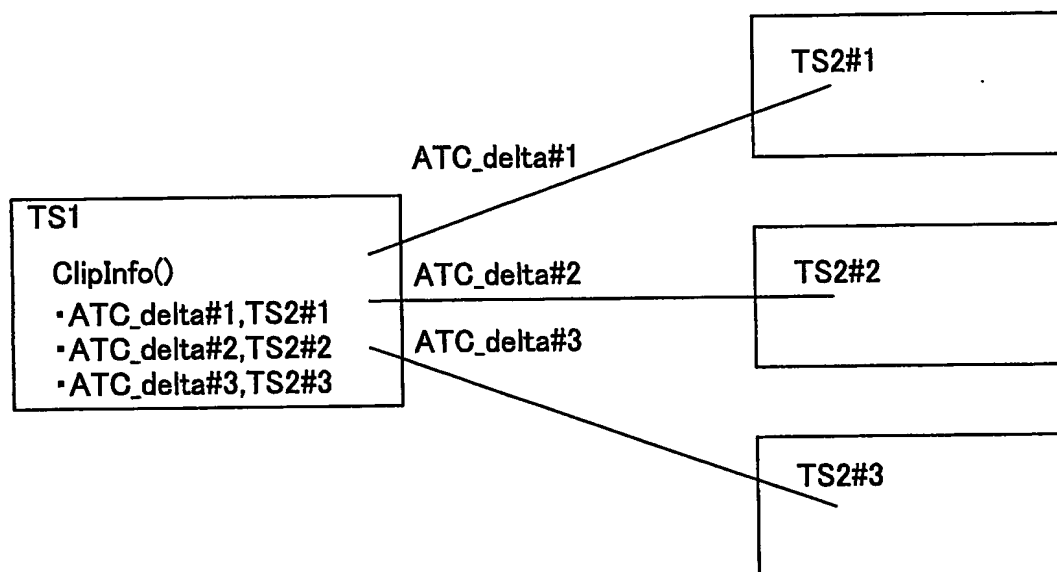


FIG.11

12/15

<u>Syntax</u>	<u>No. of bits</u>	<u>Mnemonic</u>
<u>ClipInfo()</u>		
<u>length</u>	<u>32</u>	<u>uimsbf</u>
<u>reserved for future use</u>	<u>16</u>	<u>bslbf</u>
<u>Clip stream type</u>	<u>8</u>	<u>bslbf</u>
<u>application type</u>	<u>8</u>	<u>bslbf</u>
<u>reserved for future use</u>	<u>31</u>	<u>bslbf</u>
<u>is_ATC_delta</u>	<u>1</u>	<u>bslbf</u>
<u>TS recording rate</u>	<u>32</u>	<u>uimsbf</u>
<u>number of source packets</u>	<u>32</u>	<u>uimsbf</u>
<u>reserved for future use</u>	<u>1024</u>	<u>bslbf</u>
<u>TS type info block()</u>		
<u>if (is_ATC_delta==1_b){</u>		
<u>reserved for future use</u>	<u>8</u>	<u>bslbf</u>
<u>number of ATC delta entries</u>	<u>8</u>	<u>uimsbf</u>
<u>for (i=0; i<number of ATC delta entries; i++){</u>		
<u>ATC delta[i]</u>	<u>32</u>	<u>uimsbf</u>
<u>following Clip Information file name[i]</u>	<u>8*5</u>	<u>bslbf</u>
<u>Clip codec identifier</u>	<u>8*4</u>	<u>bslbf</u>
<u>reserved for future use</u>	<u>8</u>	<u>bslbf</u>
<u>}</u>		
<u>}</u>		
<u>}</u>		

FIG.12

**FIG.13**

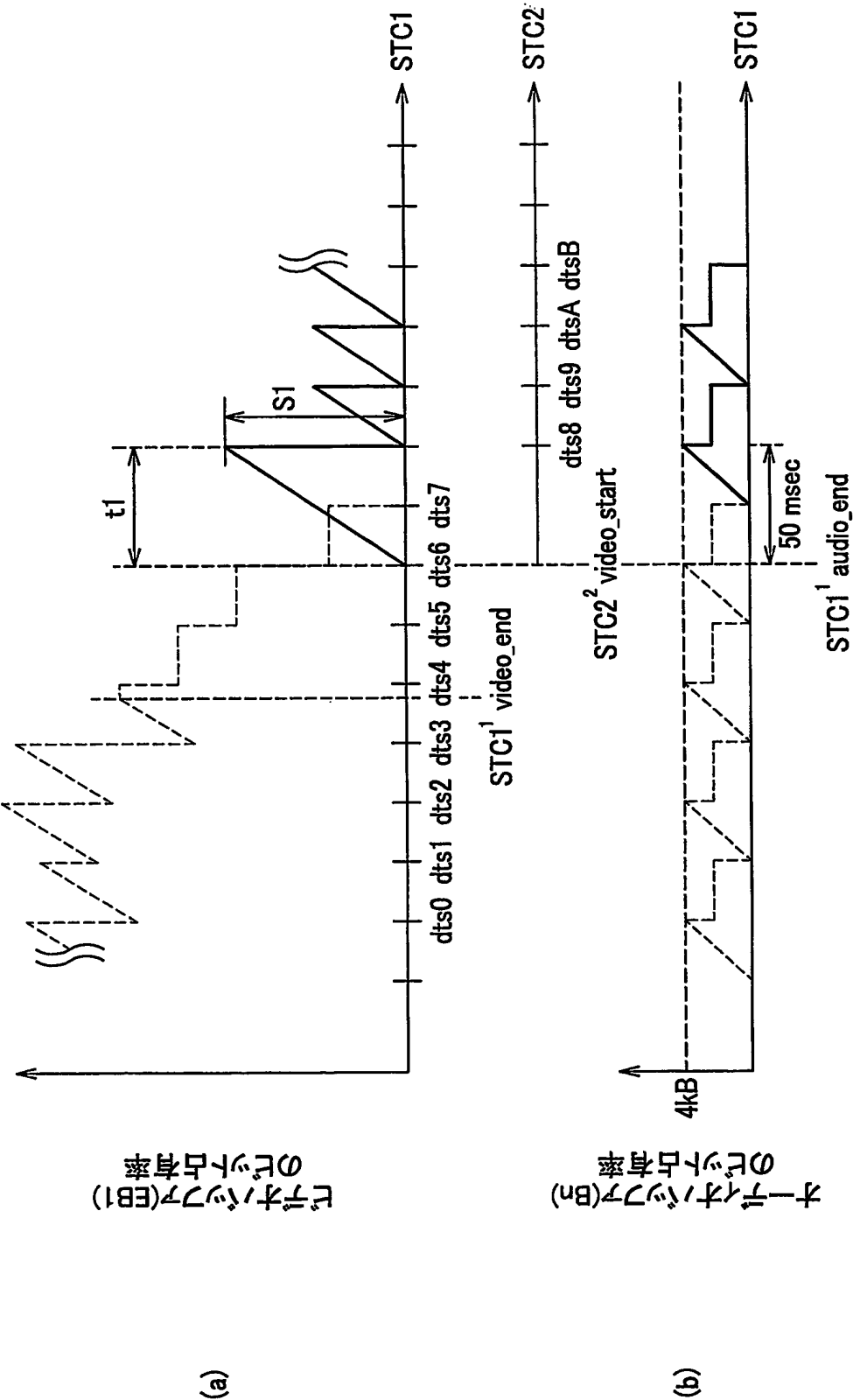


FIG.14

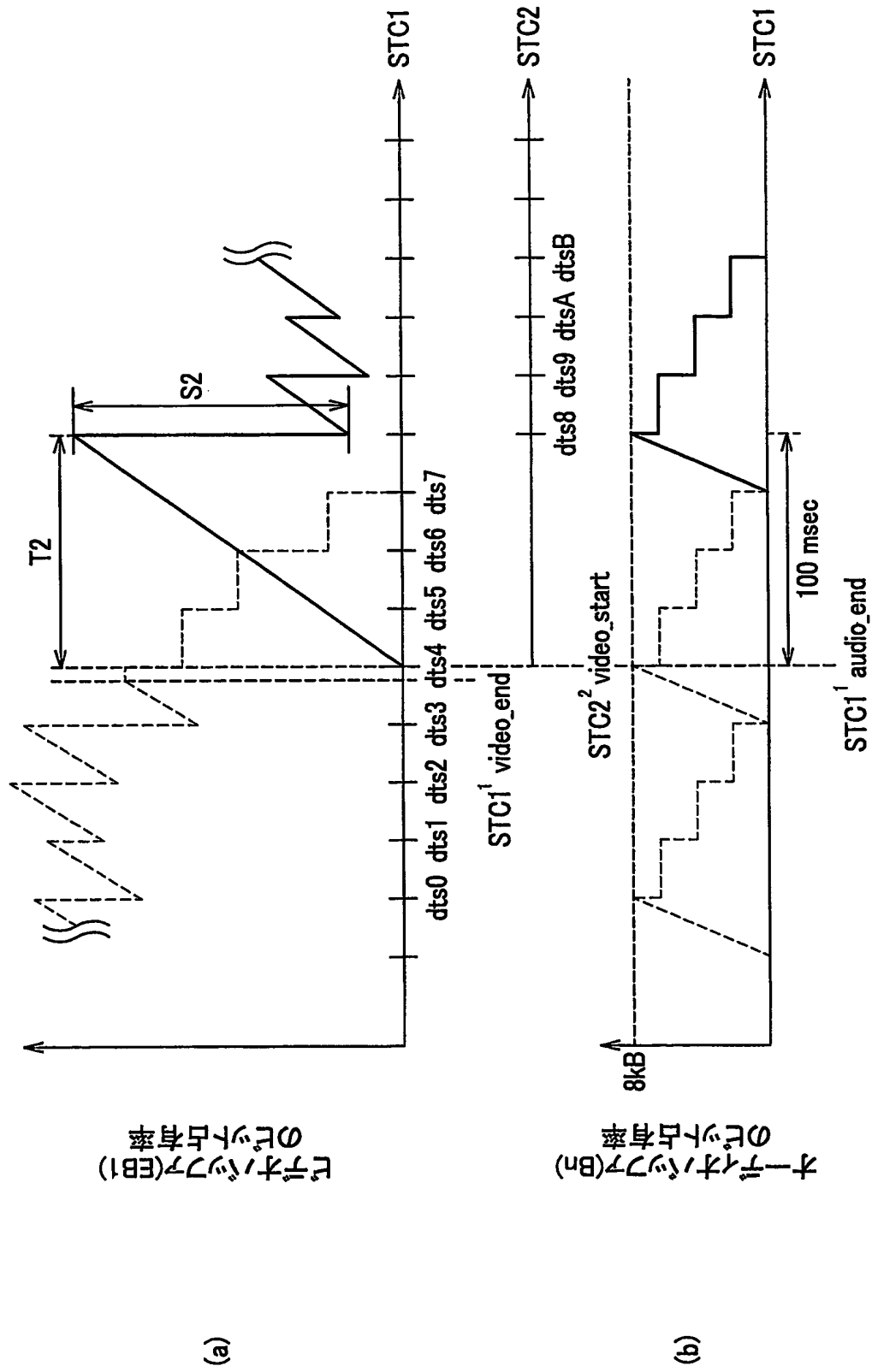


FIG.15

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/005782

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl⁷ H04N7/08, H04N7/24, H04N5/92

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ H04N7/00-7/68, H04N5/38-5/46, H04N5/76-5/956

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2004
Kokai Jitsuyo Shinan Koho	1971-2004	Toroku Jitsuyo Shinan Koho	1994-2004

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JSTPlus FILE(JOIS), IEEE Xplore

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2002-158974 A (Sony Corp.), 31 May, 2002 (31.05.02), Par. Nos. [0309] to [0327] & WO 2001/82610 A1 & EP 1198133 A1 & US 2002/145702 A1 & CN 1383679 A & KR 2002-020916 A	1-22
A	JP 2000-175152 A (Sony Corp.), 23 June, 2000 (23.06.00), Full text; all drawings (Family: none)	1-22
A	JP 2000-183957 A (Samsung Electronics Co., Ltd.), 30 June, 2000 (30.06.00), Full text; all drawings & EP 1014370 A2 & CN 1259735 A & KR 2000-47359 A	1-22

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
19 May, 2004 (19.05.04)Date of mailing of the international search report
01 June, 2004 (01.06.04)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/005782

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

The inventions of claims 1-9, 11, 13, 15, 17, 20 relate to setting of the audio buffer capacity.

The inventions of claims 10, 12, 14, 16, 18, 19, 21, 22 relate to the buffering step to the audio buffer.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int. Cl⁷ H04N 7/08
 H04N 7/24
 H04N 5/92

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int. Cl⁷ H04N 7/00 - 7/68
 H04N 5/38 - 5/46
 H04N 5/76 - 5/956

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1922-1996年
 日本国公開実用新案公報 1971-2004年
 日本国実用新案登録公報 1996-2004年
 日本国登録実用新案公報 1994-2004年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

JSTPlusファイル (JOIS)
 IEEE Explore

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	JP 2002-158974 A (ソニー株式会社) 2002. 05. 31, 段落【0309】～【0327】 & WO 2001/82610 A1 & EP 1198133 A1 & US 2002/145702 A1 & CN 1383679 A & KR 2002-020916 A	1-22

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」 特に関連のある文献ではなく、一般的技術水準を示すもの
 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
 「O」 口頭による開示、使用、展示等に言及する文献
 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
 「&」 同一パテントファミリー文献

国際調査を完了した日

19. 05. 2004

国際調査報告の発送日

01. 6. 2004

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

清水 祐樹

5P

3049

電話番号 03-3581-1101 内線 3581

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2000-175152 A (ソニー株式会社) 2000. 06. 23, 全文、全図 (ファミリーなし)	1-22
A	JP 2000-183957 A (三星電子株式会社) 2000. 06. 30, 全文、全図 & EP 1014370 A2 & CN 1259735 A & KR 2000-47359 A	1-22

第Ⅱ欄 請求の範囲の一部の調査ができないときの意見（第1ページの2の続き）

法第8条第3項（PCT17条(2)(a)）の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. ☐ 請求の範囲 _____ は、この国際調査機関が調査をすることを要しない対象に係るものである。
つまり、
2. ☐ 請求の範囲 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、
3. ☐ 請求の範囲 _____ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第Ⅲ欄 発明の単一性が欠如しているときの意見（第1ページの3の続き）

次に述べるようにこの国際出願に二以上の発明があるとこの国際調査機関は認めた。

請求の範囲1-9, 11, 13, 15, 17, 20に係る発明は、オーディオバッファ容量の設定に関するものである。

請求の範囲10, 12, 14, 16, 18, 19, 21, 22に係る発明は、オーディオバッファへのバッファリング工程に関するものである。

1. ☐ 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求の範囲について作成した。
2. ☒ 追加調査手数料を要求するまでもなく、すべての調査可能な請求の範囲について調査することができたので、追加調査手数料の納付を求めなかった。
3. ☐ 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求の範囲のみについて作成した。
4. ☐ 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求の範囲について作成した。

追加調査手数料の異議の申立てに関する注意

- ☐ 追加調査手数料の納付と共に出願人から異議申立てがあった。
☐ 追加調査手数料の納付と共に出願人から異議申立てがなかった。